

STIC-Biotech/ChemLib

146244

From: Slobodyansky, Elizabeth
Sent: Friday, February 25, 2005 9:40 PM
To: STIC-Biotech/ChemLib
Subject: 10/019,284 SEQ

Please search for case 10/019,284:

ES

SEQ ID NOs: 1 and 2 against commercail ~~and interference~~ databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner
Art Unit 1652
REM 2D83
571-272-0941
MAILBOX 2C70

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: 3/1/05
Date Completed: 3/1/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # 1 _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: 10/019,284
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 13:51:53 ; Search time 24423 Seconds
(without alignments)
11842.492 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969

Sequence: 1 agtccgtcgcacgcaccatt.....tgatatttcgcgcctgaa 5969

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_ncg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_dl: *
9: gb_pr: *
10: gb_ro: *
11: gb_stb: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5969	100.0	5969	6	BD093238	BD093238 DNA encod
C 2	5556.2	94.8	320550	1	AP005282	AP005282 Coryneb
C 3	5556.2	94.8	349136	1	BX927155	BX927155 Coryneb
C 4	5556.2	94.8	349980	6	AX127152	AX127152 Sequence
5	3687	61.8	6911	6	E11760	E11760 Base sequen
6	3687	61.8	6911	6	I26124	I26124 Sequence 4
7	1874.2	31.4	1983	6	BD165105	BD165105 Novel pol
8	1874.2	31.4	1983	6	AX122988	AX122988 Sequence
9	1457	24.4	1656	6	BD094213	BD094213 Genes for
10	1412.6	23.7	1527	6	AX069134	AX069134 Sequence
11	1412.6	23.7	1527	6	AX469820	AX469820 Sequence
12	1321.2	22.1	1342	6	AX065221	AX065221 Sequence
13	1267.8	21.2	1289	6	BD165106	BD165106 Novel pol
14	1267.8	21.2	1289	6	AX122989	AX122989 Sequence
15	1237.4	20.7	1287	6	AX065287	AX065287 Sequence
16	1237.4	20.7	1287	6	AX065289	AX065289 Sequence
17	1110.4	18.6	1152	6	BD165108	BD165108 Novel pol
18	1110.4	18.6	1152	6	AX122991	AX122991 Sequence
19	1049.8	17.6	1109	6	AX069136	AX069136 Sequence

20	762	12.8	882	6	AX065297	AX065297 Sequence
21	671.8	11.3	759	6	BD165107	BD165107 Novel pol
22	671.8	11.3	759	6	AX122990	AX122990 Sequence
23	474.2	7.9	479	6	BD094149	BD094149 Genes for
24	466.6	7.8	498	6	BD094148	BD094148 Genes for
C 25	433	7.3	343050	1	AL935252	AL935252 Lactobaci
C 26	421.6	7.1	21838	1	PDCAPOPER	L32093 Pedicoccus
C 27	421.6	7.1	21839	1	PPSUFROP	232771 P.pentoseace
28	397	6.7	5800	1	LL287015	257015 Lactococcus
29	371.8	6.2	341553	1	BX248355	BX248355 Coryneb
30	353.2	5.9	2508	1	STRSCRA	M22711 Streptococc
C 31	352.8	5.9	300330	1	AP005222	AP005222 Coryneb
C 32	351.6	5.9	9979	1	AE015011	AE015011 Streptoco
C 33	343	5.7	4750	1	AY177419	AY177419 Lactobaci
34	326.6	5.5	110000	1	AE016827_07	Continuation (8 of
35	325	5.4	2655	1	SXSCRA	X69800 S.xylosum s
C 36	323	5.4	110000	1	BX571856_25	Continuation (26 o
C 37	322.8	5.4	177911	1	AE017206	AE017206 Lactobaci
C 38	322.8	5.4	349980	6	AX926712	AX926712 Sequence
C 39	320.2	5.4	10264	1	AE004395	AE004395 Vibrio ch
40	320	5.4	2788	1	AF401046	AF401046 Lactobaci
C 41	319.8	5.4	110000	1	BX571857_24	Continuation (25 o
C 42	319.8	5.4	301250	1	AP004830	AP004830 Staphyloc
C 43	318.2	5.3	297850	1	AP003137	AP003137 Staphyloc
C 44	318.2	5.3	342350	1	AP003365	AP003365 Staphyloc
45	316.6	5.3	39425	1	AC090968	AC090968 Staphyloc

ALIGNMENTS

RESULT 1	BD093238	5969 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093238	DNA encoding sucrose PTS enzyme II.			
DEFINITION	BD093238	DNA encoding sucrose PTS enzyme II.			
ACCESSION	BD093238.1	GI:22638826			
VERSION	BD093238.1	GI:22638826			
KEYWORDS	WO 0102584-A/1.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 5969)				
AUTHORS	Izui, M., Sugimoto, M., Nakamatsu, T. and Kurahashi, O.				
TITLE	DNA encoding sucrose PTS enzyme II				
JOURNAL	Patent: WO 0102584-A 1 11-JAN-2001.				
COMMENT	OSAMU KURAHASHI				
	OS Brevibacterium lactofermentum				
	PV WO 0102584-A/1				
	PD 11-JAN-2001				
	PF 30-JUN-2000 WO 2000JP004348				
	PR 02-JUL-1999 JP 99P 189512				
	PI MASAKO IZUI, MASAKAZU SUGIMOTO, TSUYOSHI NAKAMATSU, OSAMU KURAHASHI				
FEATURES	source	1..5969			
	Location/Qualifiers				
	CC	Location/Qualifiers			
	CC	Location/Qualifiers			
	FT	Location/Qualifiers			
	CDS	Location/Qualifiers			
ORIGIN	1..5969				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
Query Match	100.0%	Score 5969	DB 6	Length 5969	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 5969	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1	AGTCCGTCGACGCGACCATTTGATGTGTGTGTCACGAGCTTGGAGGCTTTTACATCT	60		
Db	1	AGTCCGTCGACGCGACCATTTGATGTGTGTGTCACGAGCTTGGAGGCTTTTACATCT	60		

QY 61 ACGCTCCGTCGCGAGTGGGGTCAATTACGGGTGGGATCACCGCCGGTGAAGTTGCG 120
DB 61 ACGCTCCGTCGCGAGTGGGGTCAATTACGGGTGGGATCACCGCCGGTGAAGTTGCG 120
QY 121 GAACCCATGTTCTCTTGGGTGAGGAAACAGTGGGGTGAAGTTTTCGAATG 180
DB 121 GAACCCATGTTCTCTTGGGTGAGGAAACAGTGGGGTGAAGTTTTCGAATG 180
QY 181 TCGAGTTTAAAGTATGATCATCATGACTTGAAGGCTGAGTAAATTCAGTAGACCTG 240
DB 181 TCGAGTTTAAAGTATGATCATCATGACTTGAAGGCTGAGTAAATTCAGTAGACCTG 240
QY 241 CAACAGCAGGCTCAAGTCCGAAGATTAATCTAGATCCGTAGCATTAACATCATTA 300
DB 241 CAACAGCAGGCTCAAGTCCGAAGATTAATCTAGATCCGTAGCATTAACATCATTA 300
QY 301 CGTCTATGCTTGGTGAAGAAACCAATACCTGAGAAAGATGGAGAAAGTGGCAT 360
DB 301 CGTCTATGCTTGGTGAAGAAACCAATACCTGAGAAAGATGGAGAAAGTGGCAT 360
QY 361 ATCAAGAAATGAGGTCAAGCAGTTAAATAATGAGGAAAGATTTTCCCTCG 420
DB 361 ATCAAGAAATGAGGTCAAGCAGTTAAATAATGAGGAAAGATTTTCCCTCG 420
QY 421 GGGTGAATGATGCTTCTTCCAACTGAAAACGGCATCATCAAGAACTCTGAGAAC 480
DB 421 GGGTGAATGATGCTTCTTCCAACTGAAAACGGCATCATCAAGAACTCTGAGAAC 480
QY 481 CAGACCTTAAAGAGAGATTCACCCCGAACTCCCAAGATTTTCCCGTTAATG 540
DB 481 CAGACCTTAAAGAGAGATTCACCCCGAACTCCCAAGATTTTCCCGTTAATG 540
QY 541 ATCTTATATCAAGGTGAAGAGGTGGCGCTTCTTAAGGAAACGACAGACAGGCA 600
DB 541 ATCTTATATCAAGGTGAAGAGGTGGCGCTTCTTAAGGAAACGACAGACAGGCA 600
QY 601 GGAACACCGGAGTATCAACCGGAAACATGGCAACCGGATGTTGCCAAGCATGTT 660
DB 601 GGAACACCGGAGTATCAACCGGAAACATGGCAACCGGATGTTGCCAAGCATGTT 660
QY 661 CGGCGCGGCTGACGCACTGGACGCGAGGTGAAAACTTATTTCCCTGTGTGAAGAG 720
DB 661 CGGCGCGGCTGACGCACTGGACGCGAGGTGAAAACTTATTTCCCTGTGTGAAGAG 720
QY 721 TCTGTGTGCGGATTCACCTGAGAGGCTTTCATCAACGATCCGTTGTGTGCTC 780
DB 721 TCTGTGTGCGGATTCACCTGAGAGGCTTTCATCAACGATCCGTTGTGTGCTC 780
QY 781 AAAACCGGATTCATTTTCCGGCAACCCAGATCTTGGCCGGGTGATCATGCGG 840
DB 781 AAAACCGGATTCATTTTCCGGCAACCCAGATCTTGGCCGGGTGATCATGCGG 840
QY 841 GAAAAGTTGATCAATTCAGTACAGTACGCGGAACTGACATCTTTGTGAGCTTC 900
DB 841 GAAAAGTTGATCAATTCAGTACAGTACGCGGAACTGACATCTTTGTGAGCTTC 900
QY 901 TCGATCTCTGCGACGCGCAACATCATTTGCTTCCGGGCAACATGAGCAATTTG 960
DB 901 TCGATCTCTGCGACGCGCAACATCATTTGCTTCCGGGCAACATGAGCAATTTG 960
QY 961 ATACCACTACGAGCGCAATTCCTTGGCTAAAGAAAAATGTAACGCTACGCTACG 1020
DB 961 ATACCACTACGAGCGCAATTCCTTGGCTAAAGAAAAATGTAACGCTACGCTACG 1020
QY 1021 ATTGTTCATGCGATCTTCCTGATCATAGGCTCCCGGCAAGGTGGGCTTTGCT 1080
DB 1021 ATTGTTCATGCGATCTTCCTGATCATAGGCTCCCGGCAAGGTGGGCTTTGCT 1080
QY 1081 TTGCTGGGCAAGTCCGGGAGCAATATGTTGATGATCGCGGAGGCTGATTTG 1140
DB 1081 TTGCTGGGCAAGTCCGGGAGCAATATGTTGATGATCGCGGAGGCTGATTTG 1140
QY 1141 CCGATGAAAGGTGATCTAGCTGTTCAACAAAGCCTTTTTCATCAGGACGCCATG 1200

DB 1141 CCGATGAAAGGTGATCTAGCTGTTCAACAAAGCCTTTTTCATCAGGACGCCATG 1200
QY 1201 AAGCGCGGAAATGCAAGCGGTGATATTTTGGGCTTTTGAACGTACCGTCAAC 1260
DB 1201 AAGCGCGGAAATGCAAGCGGTGATATTTTGGGCTTTTGAACGTACCGTCAAC 1260
QY 1261 ATGAGTCCCGCTCTGCGGATGAGCGGCAATCCGCGGGGGGACAGCACTAGCA 1320
DB 1261 ATGAGTCCCGCTCTGCGGATGAGCGGCAATCCGCGGGGGGACAGCACTAGCA 1320
QY 1321 GTCACTGTCGACCAAGTGGGAGGATGACGCTTATGACGCGACCTTCACACT 1380
DB 1321 GTCACTGTCGACCAAGTGGGAGGATGACGCTTATGACGCGACCTTCACACT 1380
QY 1381 CAACCGTCCGCTTAAATCTCGGTCTTGGCATCAAGAAATGCTAAATCCAACTG 1440
DB 1381 CAACCGTCCGCTTAAATCTCGGTCTTGGCATCAAGAAATGCTAAATCCAACTG 1440
QY 1441 CAATTTTGTGCTTTGACTCAACCGGCGAGTGCAGAAAGGTCCATTTAGTATGAG 1500
DB 1441 CAATTTTGTGCTTTGACTCAACCGGCGAGTGCAGAAAGGTCCATTTAGTATGAG 1500
QY 1501 TACTTTAAGTACGATTAATCTATCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
DB 1501 TACTTTAAGTACGATTAATCTATCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
QY 1561 TCTGCAAGACGAGCAGAAATGCGCAAGAGTTCAGTTCATTAAGTCCCTTCCGA 1620
DB 1561 TCTGCAAGACGAGCAGAAATGCGCAAGAGTTCAGTTCATTAAGTCCCTTCCGA 1620
QY 1621 ACAAGGTTGAACCTTGGGCTTGAACAGAGTCTCAACAGTACCTTAAACCAAGAGC 1680
DB 1621 ACAAGGTTGAACCTTGGGCTTGAACAGAGTCTCAACAGTACCTTAAACCAAGAGC 1680
QY 1681 TCATTCGATGATGAGCTGGGAAAGTGTATTCAMAGATTCGAAAGGATCTTGTGG 1740
DB 1681 TCATTCGATGATGAGCTGGGAAAGTGTATTCAMAGATTCGAAAGGATCTTGTGG 1740
QY 1741 ATGAATAGTGGGACTTAAACCGGTGACATGAAGAAACGCTTAAACCAATCCGCAAG 1800
DB 1741 ATGAATAGTGGGACTTAAACCGGTGACATGAAGAAACGCTTAAACCAATCCGCAAG 1800
QY 1801 AGTTCAGTACCAATGCAATCTGTTGATGAAGAGTCTACAGCCAGATGCTGCAATC 1860
DB 1801 AGTTCAGTACCAATGCAATCTGTTGATGAAGAGTCTACAGCCAGATGCTGCAATC 1860
QY 1861 CTGATTCATGAGAGAGCTGAGATGAGGCAAAATGCTGCAAAATCCTTGAAG 1920
DB 1861 CTGATTCATGAGAGAGCTGAGATGAGGCAAAATGCTGCAAAATCCTTGAAG 1920
QY 1921 TTCAATCCTTGGCATCCGCGGAAACGCAATCGCTTTCATTTGAACATCTTCT 1980
DB 1921 TTCAATCCTTGGCATCCGCGGAAACGCAATCGCTTTCATTTGAACATCTTCT 1980
QY 1981 GTCAAGTACCAAGTTCAGGCTGCAACCTTAACTGTGAGAACAAAGCTGCAAT 2040
DB 1981 GTCAAGTACCAAGTTCAGGCTGCAACCTTAACTGTGAGAACAAAGCTGCAAT 2040
QY 2041 CTTCAACATTCGAAGAGTCCCAACCGCGCTCAACCGGATTTGGGCACTTGTG 2100
DB 2041 CTTCAACATTCGAAGAGTCCCAACCGCGCTCAACCGGATTTGGGCACTTGTG 2100
QY 2101 CCGCGGCAAAACATCTGTTGTGGCAACTGTGAAAGAAAAAGCGCATCCGCG 2160
DB 2101 CCGCGGCAAAACATCTGTTGTGGCAACTGTGAAAGAAAAAGCGCATCCGCG 2160
QY 2161 AACTGTGAAGGCCATGACTCTTCTTCCAGAGTTCATCTGTGATGCAACAAT 2220
DB 2161 AACTGTGAAGGCCATGACTCTTCTTCCAGAGTTCATCTGTGATGCAACAAT 2220
QY 2221 GCCACATCATGTTGATGAAGAGCAAGTATCAAGCTGAAAGCGCTGATCACTACG 2280

Db	2221	GCACACATCATCTGTTGATGTAAGCAGCATGTATCCAACTGTGAAAAACGCTGATCTACACG	2280
Qy	2281	TCTCATGAGCAATTTAAAGCTGCGCTAGAAACAAAAAGAAAGTACTGTGTGGGGCTATG	2340
Db	2281	TCTCATGAGCAATTTAAAGCTGCGCTAGAAACAAAAAGAAAGTACTGTGTGGGGCTATG	2340
Qy	2341	CACACAGAACTTTCCAGTTTGGCCCTTGGCCATGTGATCTCTCCGCAAGGGAGGCTC	2400
Db	2341	CACACAGAACTTTCCAGTTTGGCCCTTGGCCATGTGATCTCTCCGCAAGGGAGGCTC	2400
Qy	2401	AATGATCCCAACGGAAATGTATGTCGATGGAATACCTCCACGTCCTACTACACACGAT	2460
Db	2401	AATGATCCCAACGGAAATGTATGTCGATGGAATACCTCCACGTCCTACTACACACGAT	2460
Qy	2461	CCAGGTTTCCCCTTGCACCAAAAGCGCACCGGCTGGGCTCAACACACACCGCGTTGACC	2520
Db	2461	CCAGGTTTCCCCTTGCACCAAAAGCGCACCGGCTGGGCTCAACACACACCGCGTTGACC	2520
Qy	2521	GGACCCGACGCGATTTGCATGTGACGCACTGCGCCAGCGTCTTTTACCCGGATGCACTCTAT	2580
Db	2521	GGACCCGACGCGATTTGCATGTGACGCACTGCGCCAGCGTCTTTTACCCGGATGCACTCTAT	2580
Qy	2581	GACCTGATGTGATGCTATTCCGGTGGAGCGGATTTTATCTGACGGCAACTTAAACTTTTC	2640
Db	2581	GACCTGATGTGATGCTATTCCGGTGGAGCGGATTTTATCTGACGGCAACTTAAACTTTTC	2640
Qy	2641	TACACCCGCAACTTAAAAATTTGACGGAAAGCGCGCGCACCCCAAAACCTTGTGCAAGTC	2700
Db	2641	TACACCCGCAACTTAAAAATTTGACGGAAAGCGCGCGCACCCCAAAACCTTGTGCAAGTC	2700
Qy	2701	GAGGACCCCACTGGGCTGATGGGGCGCATTCATGCGCGTGTGCTTAAAAATCCGCTTATC	2760
Db	2701	GAGGACCCCAACTGGGCTGATGGGGCGCATTCATGCGCGTGTGCTTAAAAATCCGCTTATC	2760
Qy	2761	GACCGACCCGCGCAGGCTTTCACACCCCATTAACCGCATTCACATGATCAGCCCTGATGCT	2820
Db	2761	GACCGACCCGCGCAGGCTTTCACACCCCATTAACCGCATTCACATGATCAGCCCTGATGCT	2820
Qy	2821	GATGATTGGAACATGCTTCTTGGGGCCCAACGCGCAAAACCTCACCGGTGACGCGTTCTA	2880
Db	2821	GATGATTGGAACATGCTTCTTGGGGCCCAACGCGCAAAACCTCACCGGTGACGCGTTCTA	2880
Qy	2881	TACCGCTGCACAGATCTTGAAAACTGGGAATTTCTCCGGTGAATTCACCTTTGACCTCACT	2940
Db	2881	TACCGCTGCACAGATCTTGAAAACTGGGAATTTCTCCGGTGAATTCACCTTTGACCTCACT	2940
Qy	2941	GATGACAACTCTGTTCTGCTCTCGATTCGTTCCGATGGCTCACTGTGGAAATGCCCC	3000
Db	2941	GATGACAACTCTGTTCTGCTCTCGATTCGTTCCGATGGCTCACTGTGGAAATGCCCC	3000
Qy	3001	AACCTTTTACGCTTCGCGATGGAAGAACTGGCGAAGATCTCGACGTCGATTTTCTGT	3060
Db	3001	AACCTTTTACGCTTCGCGATGGAAGAACTGGCGAAGATCTCGACGTCGATTTTCTGT	3060
Qy	3061	CCACAAGGATTTGAGCCGAATCCACGATGAGGTTACTCACTACGCAAGCTTCGACCAAGTGC	3120
Db	3061	CCACAAGGATTTGAGCCGAATCCACGATGAGGTTACTCACTACGCAAGCTTCGACCAAGTGC	3120
Qy	3121	GGATATGTGCTGCAACAGCTTGAAAGAACGACCTTCCCGTCCGCGAGGATTTCAAGCGAG	3180
Db	3121	GGATATGTGCTGCAACAGCTTGAAAGAACGACCTTCCCGTCCGCGAGGATTTCAAGCGAG	3180
Qy	3181	CTGATTTCCGGCATGAATTTCTAGCGACCGGAGGTTGAGTAAACGGTTCTGATGCTCGG	3240
Db	3181	CTGATTTCCGGCATGAATTTCTAGCGACCGGAGGTTGAGTAAACGGTTCTGATGCTCGG	3240
Qy	3241	CTCGTGGGCTGGATGGGGCTGCGCGCAGAGTATGTCACCAACAGTTTGACAGGAAGGA	3300
Db	3241	CTCGTGGGCTGGATGGGGCTGCGCGCAGAGTATGTCACCAACAGTTTGACAGGAAGGA	3300
Qy	3301	TGGGTGCACTGCTTCACTGTGCCCCCGCAAGCTTCAATTGGCGCACACGCGATCTTACCA	3360
Db	3301	TGGGTGCACTGCTTCACTGTGCCCCCGCAAGCTTCAATTGGCGCACACGCGATCTTACCA	3360

QY	3351	NAGCTCTCTTCCCAAGCGGGAACTCCGGGGTATCAAGTCTGAAATTAAGTTTCAAGCT	3420
Db	3361	GAGCTCTTCTCCAGAGGGGGAATCGGGGTAATCAATGTGATTAAGTTTCAAGCT	3420
QY	3421	GTCCGAGTAGACATCCGAGCGAATATTTCCCTCGAGTGGATGTGTCCTTTGTCTGTG	3480
Db	3421	GTCCGAGTAGACATCCGAGCGAATATTTCCCTCGAGTGGATGTGTCCTTTGTCTGTG	3480
QY	3481	GATGTGATGGTGAATCGTCCGCTGCTGAGGTAAACCCTGGCGAATTAGTATGCGGAC	3540
Db	3481	GATGTGATGGTGAATCGTCCGCTGCTGAGGTAAACCCTGGCGAATTAGTATGCGGAC	3540
QY	3541	GATATAACGCAATTGAGATTAACGACAGGTATGACAGGTTTCATTCCCTTCCGGGC	3600
Db	3541	GATATAACGCAATTGAGATTAACGACAGGTATGACAGGTTTCATTCCCTTCCGGGC	3600
QY	3601	CTTCAAGGTGACATAATGAGATAATGATATTAAGGATCTTTTGTGCGAATTTG	3660
Db	3601	CTTCAAGGTGACATAATGAGATAATGATATTAAGGATCTTTTGTGCGAATTTG	3660
QY	3661	ACAAATCTTGCGCAAAATCCCTTATGCGACACAATAAACAAGTTTAATTTGTTAGC	3720
Db	3661	ACAAATCTTGCGCAAAATCCCTTATGCGACACAATAAACAAGTTTAATTTGTTAGC	3720
QY	3721	TTTTGAACAAACATTCAATGTCGTGAATATTTTGTTCCTCCGGTTAAGAGAAATTCAT	3780
Db	3721	TTTTGAACAAACATTCAATGTCGTGAATATTTTGTTCCTCCGGTTAAGAGAAATTCAT	3780
QY	3781	GGACCATTAAGACCTCGCGCAACGCATCTGCGCGACATTGGCGGCGAAGCAACATTGT	3840
Db	3781	GGACCATTAAGACCTCGCGCAACGCATCTGCGCGACATTGGCGGCGAAGCAACATTGT	3840
QY	3841	CGCGCGCGCACACTGTGCGCAACGCGTTTACGCTCGTGTCAAAACACCAAGATGTGGA	3900
Db	3841	CGCGCGCGCACACTGTGCGCAACGCGTTTACGCTCGTGTCAAAACACCAAGATGTGGA	3900
QY	3901	TGCGCAAAAGTGTGATGATGATTCAGATCTGAAGGACCTTTGAAATCGCGCGCATGTT	3960
Db	3901	TGCGCAAAAGTGTGATGATGATTCAGATCTGAAGGACCTTTGAAATCGCGCGCATGTT	3960
QY	3961	CCAATCATCTGTGGGCGCAAGCGCAATGTGTTTCAAAACAATCGATGACGCAAC	4020
Db	3961	CCAATCATCTGTGGGCGCAAGCGCAATGTGTTTCAAAACAATCGATGACGCAAC	4020
QY	4021	CTCCAAAGACATCGCTGTGTCCACAGACAGCTCAAAAGATTGTGTGCTAACACGCGCA	4080
Db	4021	CTCCAAAGACATCGCTGTGTCCACAGACAGCTCAAAAGATTGTGTGCTAACACGCGCA	4080
QY	4081	CTGTTTCAGCCGTGTGTGAANGTAATGGCGAACAATTTGTCCTCCGCTGATTCGAATCTT	4140
Db	4081	CTGTTTCAGCCGTGTGTGAANGTAATGGCGAACAATTTGTCCTCCGCTGATTCGAATCTT	4140
QY	4141	GTTTGGTGGCGGTCTGCTCATGCTATCAACAATGTGTGTTGCGACAGATCTGTTGCG	4200
Db	4141	GTTTGGTGGCGGTCTGCTCATGCTATCAACAATGTGTGTTGCGACAGATCTGTTGCG	4200
QY	4201	TCCGCAATCACTGTGAGATGTTCCCTCAGATCAGCGGTGTGCTGAGATCAACT	4260
Db	4201	TCCGCAATCACTGTGAGATGTTCCCTCAGATCAGCGGTGTGCTGAGATCAACT	4260
QY	4261	GATGGCAATCTGCGCGTTCCGCTTCTGGCCAGTGTGGTTTACCGCAACCAAGCG	4320
Db	4261	GATGGCAATCTGCGCGTTCCGCTTCTGGCCAGTGTGGTTTACCGCAACCAAGCG	4320
QY	4321	TTTCGATGGCATGAGTTCTGTGGCGCGGCAATTTGATGCGATGTGTTCCCAACCTT	4380
Db	4321	TTTCGATGGCATGAGTTCTGTGGCGCGGCAATTTGATGCGATGTGTTCCCAACCTT	4380
QY	4381	GGTTAACGGCTACGACGTGGCGCGCACATTAACCGCGGCGCAAAATGCTGATCTCTT	4440
Db	4381	GGTTAACGGCTACGACGTGGCGCGCACATTAACCGCGGCGCAAAATGCTGATCTCTT	4440

Dd	5521	AAACTCAAGCGCAGCACTTTAACCCGCTGAAGAAGACAGGGCGATGAATCAAGCAGG	5580
Oy	5581	GGAGCTCCTGTGTGAATTGCATATTGATTCATTAAGCCATTAGGCTGCAGGTTATAGGTAAACAC	5640
Dd	5581	GGAGCTCCTGTGTGAATTGCATATTGATTCATTAAGCCATTAGGCTGCAGGTTATAGGTAAACAC	5640
Oy	5641	GCCGATTTGGTTCGCAATTAACAAGAAACCAGGACCTGTATAACATTAAGGTTTGGGCGA	5700
Dd	5641	GCCGATTTGGTTCGCAATTAACAAGAAACCAGGACCTGTATAACATTAAGGTTTGGGCGA	5700
Oy	5701	AATTGAAGCGGAGGCCAACCCTGCTCAACGTCGCAGAAAGAAAGACGGTGCAGCAACCC	5760
Dd	5701	AATTGAAGCGGAGGCCAACCCTGCTCAACGTCGCAGAAAGAAAGACGGTGCAGCAACCC	5760
Oy	5761	ATAAGTTGAACCTTGAAGTGTTCGCACACAGTTTAGCTAGGGAGCGTGACTTAGCAT	5820
Dd	5761	ATAAGTTGAACCTTGAAGTGTTCGCACACAGTTTAGCTAGGGAGCGTGACTTAGCAT	5820
Oy	5821	CTTTGACACACCGGTAACCCGTAACGCTTGCAGATTTTTAAACCTGTTCAACCCAGTCAATGCTC	5880
Dd	5821	CTTTGACACACCGGTAACCCGTAACGCTTGCAGATTTTTAAACCTGTTCAACCCAGTCAATGCTC	5880
Oy	5881	GGTATCCTGTGTGTGTGCACACCCCGCAATCTTCAACCCCAATTGSAACATTTGCTTACG	5940
Dd	5881	GGTATCCTGTGTGTGTGCACACCCCGCAATCTTCAACCCCAATTGSAACATTTGCTTACG	5940
Oy	5941	AGTAGCGTTTGATATTTTTCGCGCGCTGAA	5969
Dd	5941	AGTAGCGTTTGATATTTTTCGCGCGCTGAA	5969
RESULT 2			
AP005282/c			
LOCUS	320550 bp	DNA	linear BCT 08-AUG-2002
DEFINITION	Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section		
VERSION	AP005282		
KEYWORDS	AP005282 BA000036		
SOURCE ORGANISM	AP005282.1 GI:21325287		
REFERENCE	1 Nakagawa,S.		
AUTHORS	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 320550)		
REFERENCE	Nakagawa,S.		
AUTHORS	Direct Submission		
TITLE	Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.		
JOURNAL	Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,		
	Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com,		
	Tel.:81-44-829-3031, Fax:81-44-813-1651)		
COMMENT	This sequence is conducted by collaboration of Kyowa Hakko Kogyo		
	Co. Ltd. And Kitasato University.		
FEATURES	location/Qualifiers		
source	1..320550		
	/organism="Corynebacterium glutamicum ATCC 13032"		
	/mol_type="genomic DNA"		
	/strain="ATCC 13032"		
	/db_xref="taxon:196627"		
	/note="ATCC 13032"		
	188..814		
gene	/gene="Cgl12516"		
	188..814		
CDS	/gene="Cgl12516"		
	/codon_start=1		
	/transl_table=11		
	/product="Hypothetical protein"		
	/protein_id="BAB99909.1"		
	/db_xref="GI:21325288"		

```

gene
CDS
/translation="MYRAFDVPLSLVLVTKRRRKHMETLAAQARTLSEKGVAP
THASFVESIAKAPILSLITLITVYNGISGCPALREOVPTDYNKINYEENIK
GLVSISEPLEHTAAQTINQRNADSDSEKVPDPGNLVYLQONIPYANANADTTVDR
FLNSPDHVKLLANDYEATGVAAIKGDHAWIVERTVAPADSVESTE"
811..2151
/gene="Cg12517"
811..2151
/gene="Cg12517"
/codon_start=1
/transl_table=11
/product="Nicotinic acid phosphoribosyltransferase"
/protein_id="BAB99910.1"
/db_xref="GI:21325289"
/translation="MNTNPEEFSNRSSTALLTDKYEELTMQLADGNERSTFEV
SRRLNERRIVAGTAVKAIKIRPFVFTBEQLADLDPLDDTLTLKRYRTGVGDG
YREGIYFEPQSLTVRGTFACVILETVLSEIMNDSVSAARWYVAADGRPIIE
MGSRTHEYSATYASRAAYLAGSTTSLEAARYGIPASGTSAAHMTLHINDGTP
NEAARFKQVESLGVDTLLVDYDITOGVATVIEVAGDLGVRIDSGDLVARKY
RKOLDLNAHNTKIVVSSDLDEFAIAGLGEVDVGVTGVSATVAGTAVLAKIV
EVAGHPVAKRSNKSXGGKKAIVTHRSQAIKEIYVPPAAEPDQKDLTSLTI
PLMRGEIYVPGIPTLEDSRAYLAKOLVSLPMEGLASRDEPVLHTRFVGFPPAA"
2438..2746
/gene="Cg12518"
2438..2746
/gene="Cg12518"
/codon_start=1
/transl_table=11
/product="Hypochemical protein"
/protein_id="BAB99911.1"
/db_xref="GI:21325290"
/translation="MGLRLNLRGPRPOKQITWKLPAONRPOAPLSKOSNOTVNRCPA
WLSLSTPIWLDLAQCLRLNHPQIPPOIISMLEPKSCLLPLKRMVAVVLDKQMLRL
"
2866..4608
/gene="Cg12519"
2866..4608
/gene="Cg12519"
/codon_start=1
/transl_table=11
/product="Rad3-related DNA helicases"
/protein_id="BAB99912.1"
/db_xref="GI:21325291"
/translation="MSTATIALROLQVNRDLPRLVDALTEPLMERPTPAIMKGRNVL
CMNKYAROEINOBDALIEQDISLWLGKIVLNMWANEETEGDDDDIDRCYDILAMK
QVSVTARBIGTASRCRPHGEDCPAETARGAKADAVVTHALALDALSIVSEPHD
VVVIDEAHELDERTAVASAEITVNSLNLAAARAKLSDSKREERVOBIAGLETLQ
TMQGRNMDMEGSKGTLVALKDALMALAQIAGAPEGEAANDPERFAERONLSHLM
EIHDAVNRILIEVFAEDPSKQYDVVNNHNDRGRSLNVAPLSVAGLHEKLFASNTV
VLASATLITIGNFNMAAASWGLPKGSNMDAGCTPDPAKSGIILYTARHLLPDGDL
PBETIDEIYELITAGGRTGLFSSKRAAEQTKMRRLRPPDVLCQDDNTALVYK
FSDSENTCLFGTLTMQGVDPGRSLVLDRIFFPPDDPLQARKEADAEGRNG
FMEVATHAALLMAQAGRLLRHVGDGVVALDHRLLTKRYGGLRFSMPRFWEITN
PETVRAALKRLVTK"
complement(4618..5511)
/gene="Cg12520"
complement(4618..5511)
/gene="Cg12520"
/codon_start=1
/transl_table=11
/product="Hypochemical membrane protein"
/protein_id="BAB99913.1"
/db_xref="GI:21325292"
/translation="WTSFYGSQYPGDDNNNNNSQFGNPSGEQNTQPYGAPGYQO
PFDGFNAYSSPIPREVPQPSMQEAKWMSFDGTVFGAMKPTATQAWVLSALIF
AVLVLNFMAMILPMVSVLAATSSGDSALIAATGTSFFGFWLMTVLAIFSVSLIK
YNNARVVRGEQITIOSFFKMGKLGALGIYLLINIVIFIGMILLIPGIIAAVVLIF
AVPVAFLRDAIDGAFSASWKAIVKNGQVILLBLAIFALSFLGSNVIIGMLVTTPL
TFLLAYAFQTSQSGPIMORQ"
complement(5587..6312)
/gene="Cg12521"
complement(5587..6312)
/gene="Cg12521"
/notice="PF01981:Domain of unknown function UPF0099"

gene
CDS
/codon_start=1
/transl_table=11
/product="Uncharacterized ACR"
/protein_id="BAB99914.1"
/db_xref="GI:21325293"
/translation="MPKDSLDLHLLKARASDGPSEDPSPSTVOAMQIALHIPEQ
NPPRTDVLAAARSVKLCIDERVSTPDPRALIERMYGHLIRKVSRRANVADR
ODLRGVYEDDAAKRATRLPSAVYDVPADIKKQISGSELPIDENPNIPRNDVYID
ESLKNITLGRALAAQVGHASMLLAHQPFVMEVQMEADALHVRILPSEEFLLISPG
AVPVRDGGFTVAPNTVVAIP"
complement(6305..7324)
/gene="Cg12522"
complement(6305..7324)
/gene="Cg12522"
/codon_start=1
/transl_table=11
/product="Phosphoserine phosphatase"
/protein_id="BAB99915.1"
/db_xref="GI:21325294"
/translation="MHGQSVVVELOETVQSSRPRESHVVVVVGDPVADLISRIGOTL
ADVANDITIRGISDPYTGLEIKTVPDVSPGGEAMRKALALITSLANDIATERS
GLIRSKRLVCPDCDSTIITGEVIMLAHAKGEVAVATERARNGELDEPRESLER
VKALAGLASVYIDEVAAAEILTPEKRTIRTLNRGYOTAVVSGFIVLGLAELE
LDVYRANTLEIVDGLTNGVNGKIYDRAAKAFLEFPADSGLKMYQTVVAGDGNADI
DMLSAAGVAFNAKPALEKAIADTSVNHPLDEVLIHNGISRDEIDLQDEGCTHRY
PLTNA"
complement(7811..9565)
/gene="Cg12523"
complement(7811..9565)
/gene="Cg12523"
/EC_number="1.9.3.1"
/notice="PF00115:Cytochrome C and Quinol oxidase polypeptide
I"
/codon_start=1
/transl_table=11
/product="Heme/copper-type cytochrome/quinol oxidases,
subunit 1"
/protein_id="BAB99916.1"
/db_xref="GI:21325295"
/translation="WTAAPRVVDGVAPQREPPGCHARKSKAMLTMTTHKQIGIM
YIISFSFFELGGLALIRALEFTPGIQLSNEQFNQLFTMKGVMILLTGTPIWVG
FANVYLPQIADPAPVAPRLNAFGWITTVGVAMLTGFLTPGGAADGWTMYSLSD
AIHFPGLSGDMWITGVGATIGISVASAINMTLLICLRAPGTMRMPDIFWNIIFVS
VLLALIFPRLAALAGVYXDRKLGHLVDPRANGSLMQLHFWPFGHEVVYALPFI
GIVERILIPVPRKMPGVYGLIPLTSLGALSMAVMAMHMTVGAVLLPSPFMTFLI
SVPTGVKFFRNVTGMKGHITWETPMISVGPATFLFGLTGIMLAPPLDPLILADS
YFLIAHFYTLFETVVPASCAVYFWFPMKMGMMDEKLIHFWLTLSPGHGTFLIQ
HVMGMGMPPRYADVLDSDGFTLIQNSTVSPFLGLSVIPIFIMVFSWYGEGLVTV
DDPWGNSLSEWATSCPPPRNPFASLPRIRESRPAFELIYPHMIERMAEAHTGHDD
INAPELGTAPALASDSR"
9701..10009
/gene="Cg12524"
9701..10009
/gene="Cg12524"
/codon_start=1
/transl_table=11
/product="Hypochemical protein"
/protein_id="BAB99917.1"
/db_xref="GI:21325296"
/translation="MPKNVQVTHVFAARSHLMAVTPNORCEGVNLTDISPGSP
FTVVPCHRFHNSPRGSIEMLRPFQGRFLIKRVVPPRGPGAWMSPTPICASEKPA
P"
complement(10086..11090)
/gene="Cg12525"
complement(10086..11090)
/gene="Cg12525"
/notice="PF00268:Ribonucleotide reductase, small chain"
/transl_table=11

```


Db 150476 CCGATCATAGAGACAGCTGACAGTATGAGGCAAAAGATGCTGACAGATCCGTTGAT 150417
Qy 1920 GTTCAATCTTTGGCATCGGCGGAAACGG-CACATGCTTTTCAATGAAACATCATCTTCT 1978
Db 150416 GTTCAATCTTTGGCATCGGCGGAAACGG-CACATGCTTTTCAATGAAACATCATCTTCT 150357
Qy 1979 CTGTACAGACTGACAAAGGTCCAGGCGGTGACCCCTTAAACGTGTGAGGACAAACGTCTGA 2038
Db 150356 CTGTACAGACTGACAAAGGTCCAGGCGGTGACCCCTTAAACGTGTGAGGACAAACGTCTGA 150297
Qy 2039 TTCTTCAACACATCGAAGAGGTCCCAACCCAGCGCTGACCCAGGGTTTGGGCACTTTG 2098
Db 150296 TTCTTCAACACATCGAAGAGGTCCCAACCCAGCGCTGACCCAGGGTTTGGGCACTTTG 150237
Qy 2099 TCCCGCGGCAAAACATCGTGTGTGTGCACTGTGAAAGAAAGCCGACCGCATCCG 2158
Db 150236 TCCCGCGGCAAAACATCGTGTGTGTGCACTGTGAAAGAAAGCCGACCGCATCCG 150177
Qy 2159 GGAACGTGTGAAAGGCCAGTGACTGTTCTTGCCCAAGTTTCAATCCTGTATGATGACAAAC 2218
Db 150176 GGAACGTGTGAAAGGCCAGTGACTGTTCTTGCCCAAGTTTCAATCCTGTATGATGACAAAC 150117
Qy 2219 ATGCCACATCATCTGTGTGATGAGAGAGATATCCAGCTGGAAGAACCTGATCACTAC 2278
Db 150116 ATGCCACATCATCTGTGTGATGAGAGAGATATCCAGCTGGAAGAACCTGATCACTAC 150057
Qy 2279 CGTCTCATGAGCAATTTAAAGCTGCGCTAGAAAACAAAAGAAAGTACTGTGTGGGCTA 2338
Db 150056 CGTCTCATGAGCAATTTAAAGCTGCGCTAGAAAACAAAAGAAAGTACTGTGTGGGCTA 149997
Qy 2339 TGCACACAAACCTTCCAGTTTTCGCGCGCGGCAATGATGATCCTCTCCGAGGGAGAGC 2398
Db 149996 TGCACACAAACCTTCCAGTTTTCGCGCGCGGCAATGATGATCCTCTCCGAGGGAGAGC 149937
Qy 2399 TCAATGATCCCAACGGAATGTAGCTGATGAGATACCTTCAACGTCTACTACACAGAC 2458
Db 149936 TCAATGATCCCAACGGAATGTAGCTGATGAGATACCTTCAACGTCTACTACACAGAC 149877
Qy 2459 ATCCAGGTTTCCCTTTCGACCAAGAGCGACCGGCTGGGCTCAACACCAACGCGGTTGA 2518
Db 149876 ATCCAGGTTTCCCTTTCGACCAAGAGCGACCGGCTGGGCTCAACACCAACGCGGTTGA 149817
Qy 2519 CCGGACCGGAGGATTTGAGTGGAGGACGACCTGCGCGACGCTCTTTCACCGGATGATCCT 2578
Db 149816 CCGGACCGGAGGATTTGAGTGGAGGACGACCTGCGCGACGCTCTTTCACCGGATGATCCT 149757
Qy 2579 ATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2638
Db 149756 ATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 149697
Qy 2639 TCTTACACCGGCAACCTTAAATTTGACGAAAGCGCGCGCAACCCAAAACCTTGTGAGAG 2698
Db 149696 TCTTACACCGGCAACCTTAAATTTGACGAAAGCGCGCGCAACCCAAAACCTTGTGAGAG 149637
Qy 2699 TCGAGAGCCCAACTGGGCTGATGAGGAGGATTCATGCGCGTTCCTTAAATTCGCTTA 2758
Db 149636 TCGAGAGCCCAACTGGGCTGATGAGGAGGATTCATGCGCGTTCCTTAAATTCGCTTA 149577
Qy 2759 TCGACGGAACCGGCAACCGGCTTTCACACCCCATTTACCGGATCCCATGATCAAGCCCTGAG 2818
Db 149576 TCGACGGAACCGGCAACCGGCTTTCACACCCCATTTACCGGATCCCATGATCAAGCCCTGAG 149517
Qy 2819 GTGATGTTGGAACATGTTTCTTGGGCGCAACGCGAAGAAACCTCACCGGTGACGCGTTT 2878
Db 149516 GTGATGTTGGAAGAAATGTTTCTTGGGCGCAACGCGAAGAAACCTCACCGGTGACGCGTTT 149457
Qy 2879 TATACCGCTGACAGATCTTTGAAAACTGGAATTTCTCCGTTGAAATCACTTTGACCTGA 2938
Db 149456 TATACCGCTGACAGATCTTTGAAAACTGGAATTTCTCCGTTGAAATCACTTTGACCTGA 149397
Qy 2939 GTGATGCAACACTGTTCTGCTCTGATCTCGTTCCGATGAGCTACATGTGGAAATGCC 2998
Db 149396 GTGATGCAACACTGTTCTGCTCTGATCTCGTTCCGATGAGCTACATGTGGAAATGCC 149337

Qy 2999 CCAACCTTTTTCAGCTTGGCCGATGAAAGAACTGGGAAAGATCTCGACGCTGATTTTCT 3058
Db 149336 CCAACCTTTTTCAGCTTGGCCGATGAAAGAACTGGGAAAGATCTCGACGCTGATTTTCT 149277
Qy 3059 GTTCAAGAGATTTGAGCCGAAATCCACGATGAGGTTACTCATACCAAGCTCTGACAGT 3118
Db 149276 GTTCAAGAGATTTGAGCCGAAATCCACGATGAGGTTACTCATACCAAGCTCTGACAGT 149217
Qy 3119 GCGGATATGTCGTGACAAAGCTTGAAGAACGACCTTTCGCGTCTTTCGAGGATTTCAAGC 3178
Db 149216 GCGGATATGTCGTGCGCAAGCTTGAAGAACGACCTTTCGCGTCTTTCGAGGATTTCAAGC 149157
Qy 3179 ACTGGAATTTCCGCAATGATTTCTACGACCCGACAGTTGCAATTAACGTTTGTATGCTT 3238
Db 149156 ACTGGAATTTCCGCAATGATTTCTACGACCCGACAGTTGCAATTAACGTTTGTATGCTT 149097
Qy 3239 GCGTGGTGGCTGATGAGGCGCTGCGCGGCAAGATGATCAACCAAGTTGACAGAGAG 3298
Db 149096 GCGTGGTGGCTGATGAGGCGCTGCGCGGCAAGATGATCAACCAAGTTGACAGAGAG 149037
Qy 3299 GATGGGTGACATGCTGACCTGACTGTGCCCGCAAGCTTCAATTTGGCAACACGCGATCTAC 3358
Db 149036 GATGGGTGACATGCTGACCTGACTGTGCCCGCAAGCTTCAATTTGGCAACACGCGATCTAC 148977
Qy 3359 AAGAGCTCTTTCCTCCAGAGGGGAGTGGGGGTTATCAGATCTGTATTAAGTTCTGAAAC 3418
Db 148976 AAGAGCTCTTTCCTCCAGAGGGGAGTCAAGGGGTTATCAGATCTGTATTAAGTTCTGAAAC 148917
Qy 3419 CTGTCCGATGACATCCGAGGCAATATTTCCCTCGAGTGGAGTGTGTCGCTTGTCTG 3478
Db 148916 CTGTCCGATGACATCCGAGGCAATATTTCCCTCGAGTGGAGTGTGTCGCTTGTCTG 148857
Qy 3479 TGGATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3538
Db 148856 TGGATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 148797
Qy 3539 ACGATATACAGCCATTGAGATATCTGCAAGGATGATGACAGGTTTCACTTCCGCTTCCG 3598
Db 148796 ACGATATACAGCCATTGAGATATCTGCAAGGATGATGACAGGTTTCACTTCCGCTTCCG 148738
Qy 3599 GCCTTCAAGGATGACATTTAGAGATTAAGATATTAAGAGGCTCTTGTGGCGAAT 3658
Db 148737 GCCTTCAAGGATGACATTTAGAGATTAAGATATTAAGAGGCTCTTGTGGCGAAT 148678
Qy 3659 GTACAAATACCTTTCGAAATTCCTTTCGATTCGACACAAATTAACAGGTTTATATTTTGA 3718
Db 148677 GTACAAATACCTTTCGAAATTCCTTTCGATTCGACACAAATTAACAGGTTTATATTTTGA 148618
Qy 3719 GCTTTTGAACAAACATTCATGCTGATATTTTCTTCCGCTTTCGAGGATTAAGAGAAATTC 3778
Db 148617 GCTTTTGAACAAACATTCATGCTGATATTTTCTTCCGCTTTCGAGGATTAAGAGAAATTC 148558
Qy 3779 ATGGAACATTAAGACCTGCGGCAAGCATCTGCGGCAATTTGGGCGGCAACATTT 3838
Db 148557 ATGGAACATTAAGACCTGCGGCAAGCATCTGCGGCAATTTGGGCGGCAACATTT 148498
Qy 3839 GTGCGCGCGGCAACATGTTGCAACGCGTTTTCGCTGCTGCTCAAGACACCAAGATGTC 3898
Db 148497 GTGCGCGCGGCAACATGTTGCAACGCGTTTTCGCTGCTGCTCAAGACACCAAGATGTC 148438
Qy 3899 GATCCGCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3958
Db 148377 GATCCGCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 148318
Qy 3959 TTCCAGATCATGTCGGGCGCAGGCGATGATGATGATGATGATGATGATGATGATGATGAT 4018
Db 148377 TTCCAGATCATGTCGGGCGCAGGCGATGATGATGATGATGATGATGATGATGATGATGAT 148318
Qy 4019 ACCTCAAGACATGCTGTGTTCACAGAGAGCTCAAGAGTGTGTGCTTAAACAGCC 4078
Db 148317 ACCTCAAGACATGCTGTGTTCACAGAGAGCTCAAGAGTGTGTGCTTAAACAGCC 148258

QY 4079 AACTGTTCAAGCCGTGCTGTGAAGATTTGCGGACATTTTCTGCTCCGCTAATTCATC 4138
 DB 148257 AACTGTTCAAGCCGTGCTGTGAAGATTTGCGGACATTTTCTGCTCCGCTAATTCATC 148198
 QY 4139 TTGGTTGGTGGCGGTCTCTCAATGGCTATCAACATATGTTGGTGGCGGACGATCTGTC 4198
 DB 148197 TTGGTTGGTGGCGGTCTCTCAATGGCTATCAACATATGTTGGTGGCGGACGATCTGTC 148138
 QY 4199 GGTCCGCAATCACTGTGTGAAGATGTTCCCTCAGATCAGCGGTGTGTGAGATGATCAAC 4258
 DB 148137 GGTCCGCAATCACTGTGTGAAGATGTTCCCTCAGATCAGCGGTGTGTGAGATGATCAAC 148078
 QY 4259 CTGATGCAATCTGCGCGCTTCCGCTTCTTCCAGTGTGTGTGTTTCAACGCAACCAAG 4318
 DB 148077 CTGATGCAATCTGCGCGCTTCCGCTTCTTCCAGTGTGTGTGTTTCAACGCAACCAAG 148018
 QY 4319 CGTTTCGGTGGCAATGAGTTCCTGGGCGCGGCAATGGATGATGATGTTTCCCAAC 4378
 DB 148017 CGTTTCGGTGGCAATGAGTTCCTGGGCGCGGCAATGGATGATGATGTTTCCCAAC 147958
 QY 4379 CTGTTTAAACGCTAAGCACTGCGCGCAACATGACCGCGGCGAAATGCAATGTGTCC 4438
 DB 147957 TTGGTAAACGCTAAGCACTGCGCGCAACATGACCGCGGCGAAATGCAATGTGTCC 147898
 QY 4439 CTGTTTGGTGGATGTTGCTCAAGCTGTATCAAGGGCAACCTGCTTCTGTGTGTGTC 4498
 DB 147897 CTGTTTGGTGGATGTTGCTCAAGCTGTATCAAGGGCAACCTGCTTCTGTGTGTGTC 147838
 QY 4499 GTCCTGGAATCTGGGCAACATGCAAGAAATCTGCAACAAAGCACTATGGGCACTCA 4558
 DB 147837 GTCCTGGAATCTGGGCAACATGCAAGAAATCTGCAACAAAGCACTATGGGCACTCA 147778
 QY 4559 GACTTCTGATCAACCCAGTGTGAATCTGCTGTCAACCGGCTCTTACGTTACTGCT 4618
 DB 147777 GACTTCTGATCAACCCAGTGTGAATCTGCTGTCAACCGGCTCTTACGTTACTGCT 147718
 QY 4619 ATTGGTCAACGATGGCTGGGTGGGTGACTGCTGGCAACAGCTGTGAGGAGCTCAT 4678
 DB 147717 ATTGGTCAACGATGGCTGGGTGGGTGACTGCTGGCAACAGCTGTGAGGAGCTCAT 147658
 QY 4679 GATTTCCGATGCTCAGTCGCGGCTGCTGCTTTCGCTGTGCTACTCAACCAATCGTTATC 4738
 DB 147657 GATTTCCGATGCTCAGTCGCGGCTGCTGCTTTCGCTGTGCTACTCAACCAATCGTTATC 147598
 QY 4739 ACTGCTGCAACCAATGCTTCCCGCAATTTAGCTGTGAGCTTTCACCAAGGTGTGATCC 4798
 DB 147597 ACTGCTGCAACCAATGCTTCCCGCAATTTAGCTGTGAGCTTTCACCAAGGTGTGATCC 147538
 QY 4799 TTCAATCTGCAACCGCATTCATGGCCATATCGCGCAGGAGTGTGAGCATGTTGGCAGTG 4858
 DB 147537 TTCAATCTGCAACCGCATTCATGGCCATATCGCGCAGGAGTGTGAGCATGTTGGCAGTG 147478
 QY 4859 TTCTTCTGAGCAAGATGAAAAAGCTCAAGGGCCTTTCAGAGTGTCTCCGCT 4918
 DB 147477 TTCTTCTGAGCAAGATGAAAAAGCTCAAGGGCCTTTCAGAGTGTCTCCGCT 147418
 QY 4919 GTTCTTGTGATTAACAGACCTCGCATCTTCGCTGTGAACCTTCGCTGCGCTGCGCTTC 4978
 DB 147417 GTTCTTGTGATTAACAGACCTCGCATCTTCGCTGTGAACCTTCGCTGCGCTGCGCTTC 147358
 QY 4979 TACATTTGATCGGTAACCGAGCTATCGGTGGGCGCTTATGATGCACTCTTATATATCAAG 5038
 DB 147357 TACATTTGATCGGTAACCGAGCTATCGGTGGGCGCTTATGATGCACTCTTATATATCAAG 147298
 QY 5039 GCAGTTGCGTTGGCGCTGCAAGTTTCTTGGGTGTTGTTTCTAATGATGCTCAAGATATG 5098
 DB 147297 GCAGTTGCGTTGGCGCTGCAAGTTTCTTGGGTGTTGTTTCTAATGATGCTCAAGATATG 147238
 QY 5099 GTCAATGTTCTGTGTTGGCGCGGATGATCTTGTATGCAATGCGGCCCAAGCATGCT 5158
 DB 147237 GTCAATGTTCTGTGTTGGCGCGGATGATCTTGTATGCAATGCGGCCCAAGCATGCT 147178
 QY 5159 TATGCGCTTACTGTGTTGCGCGCAACGGCAGCATGATGATCAAGATGCAACCGCTGTCTCA 5218

DB 147177 TATGCGCTTACTGTGTTGCGCGCAACGGCATTTATCAATGCAACCGCTGTCTCA 147118
 QY 5219 GTGCGCTCAGGAACGACCAAAAGCGGAAGCAAGACCCGAGATTTTCAAGATTC 5278
 DB 147117 GTGCGCTCAGGAACGACCAAAAGCGGAAGCAAGACCCGAGATTTTCAAGATTC 147058
 QY 5279 ACCATATCCAGGACCTTTGACCGGTGAAGCTATCGCACTGAGACCGTCAAGATGCC 5338
 DB 147057 ACCATATCCAGGACCTTTGACCGGTGAAGCTATTCACCTGAGACCGTCAAGATGCC 146998
 QY 5339 ATGTTTGGCAGCGGAAAGCTTGGCTCAGGTGTGCAATGCTGCCCAACCAAGGCGCAGCTG 5398
 DB 146997 ATGTTTGGCAGCGGAAAGCTTGGCTCAGGTGTGCAATGCTGCCCAACCAAGGCGCAGT 146938
 QY 5399 GTTTTCAACGATGACCGGAAGATGAGTGTGTGCTTCCATCTGTCAACGCTTTCGACGTC 5458
 DB 146937 GTTTTCAACGATGAGTGAAGATGAGTGTGTGCTTCCATCTGTGCAACGCTTTCGACGTT 146878
 QY 5459 CGCACTAAGCTGAGATGTTCCATATGTGATATCTTGATGCAATTTGTTTGACACCC 5518
 DB 146877 CGCACTAAGCTGAGATGTTCCATATGTGATATCTTGATGCAATTTGTTTGACACCA 146818
 QY 5519 GTAAACCTCAACGGCAGCACTTTAACCCGCTGAAGAAAGCAGGCGATGAAGTCAAGCA 5578
 DB 146817 GTAAACCTCAACGGCAGCACTTTAACCCGCTGAAGAAAGCAGGCGATGAAGTCAAGCA 146758
 QY 5579 GGGAGCTGCTGTGTGAATTCGATATGATGATCCATTAAGCTGCAAGTTATGAGTAAAC 5638
 DB 146757 GGGAGCTGCTGTGTGAATTCGATATGATGATCCATTAAGCTGCAAGTTATGAGTAAAC 146698
 QY 5639 ACGCGATGTTGTTGTTGAAATTAACAAGAAACCGGACCTGTAAACATTAAGTTTGGGC 5698
 DB 146697 ACGCGATGTTGTTGTTGAAATTAACAAGAAACCGGACCTGTAAACATTAAGTTTGGGC 146638
 QY 5699 GAAATGGAAGCGGAGGCAACCTGCTCAACGCTGCAAGAAAGAAAGCGGTCCAGCAACA 5758
 DB 146637 GAAATGGAAGCGGAGGCAACCTGCTCAACGCTGCAAGAAAGAAAGCGGTCCAGCAACA 146578
 QY 5759 CCATATGTTGAACCTTTGAGTGTTCGCAACAGGTTAGCTAGGGGACGTGACTCTACGC 5818
 DB 146577 CCATATGTTGAACCTTTGAGTGTTCGCAACAGGTTAGCTAGGGGACGTGACTCTACGC 146518
 QY 5819 ATCTTTGACACCGGATCCGATGAGCTTTCGAGATTTTAAACCTGTTCACCAAGCTATGCC 5878
 DB 146517 ATCTTTGACACCGGATCCGATGAGCTTTCGAGATTTTAAACCTGTTCACCAAGCTATGCC 146458
 QY 5879 TCGGTATCCTGTGTGTGTCACCCCGCATCTTTCACCCCATTTGAGCATGTTGTTCA 5938
 DB 146457 TCGGTATCCTGTGTGTGTCACCCCGCATCTTTCACCCCATTTGAGCATGTTGTTCA 146398
 QY 5939 GCAATGAGCTTGTGATTTTTCGCGCGCTG 5967
 DB 146397 GCAATGAGCTTGTGATTTTTCGCGCGCTG 146369

RESULT 3

BX927155/c 349136 bp DNA linear BCT 10-JUN-2004

LOCUS Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5,

DEFINITION complete genome; segment 8/10.

ACCESSION BX927155 BX927147

VERSION BX927155.1 GI:41326514

KEYWORDS complete genome.

SOURCE Corynebacterium glutamicum ATCC 13032

ORGANISM Corynebacterium glutamicum ATCC 13032

Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;

Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1 (bases 1 to 349136)

Kalinowski, J., Bathe, B., Barfels, D., Bischoff, N., Bort, M.,

Burkowski, A., Dusch, N., Eggeling, L., Elkmann, B., J., Gajalac, L.,

Goemann, A., Hartmann, M., Huttmacher, K., Kramer, R., Linke, B.,

McHardy, A.C., Meyer, F., Mockel, B., Pfeifferle, W., Puhler, A.,

TITLE	JOURNAL	COMMENT	FEATURES
<p>Key, D.A., Ruckert, C., Rupp, O., Sahn, H., Wendisch, V.F., Wiegand, I. and Tsuchi, A.</p> <p>The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins</p> <p>J. Biotechnol. 104 (1-3), 5-25 (2003)</p>	<p>JOURNAL</p> <p>12948626</p> <p>2 (bases 1 to 349136)</p> <p>Kalinowski, J.</p> <p>Direct Submission</p> <p>Submitted (21-JAN-2004) Joern Kalinowski, Institut fuer Genomforschung, Universitaet Bielefeld, Universitaetsstrasse 25, 33615 Bielefeld, Germany</p> <p>E-mail: Joern.Kalinowski@cbitec.uni-bielefeld.de</p> <p>This sequence was accomplished by collaboration between Degussa AG and Bielefeld University.</p> <p>join(BX927148.1:1.348071, BX927149.1:1.3490887, BX927150.1:1.348475, BX927151.1:1.349459, BX927152.1:1.349799, BX927153.1:1.349584, BX927154.1:1.349575, BX927155.1:1.349136, BX927156.1:1.349115, BX927157.1:1.340057).</p> <p>location:Qualifiers</p>	<p>1.349136</p> <p>/organism="Corynebacterium glutamicum ATCC 13032"</p> <p>/mol_type="genomic DNA"</p> <p>/strain="DSM 20300 = ATCC 13032"</p> <p>/db_xref="taxon:196627"</p> <p>/note="1S fingerprint type: 4-5"</p> <p>complement(67..729)</p> <p>/locus_tag="cg2562"</p> <p>complement(67..729)</p> <p>/locus_tag="cg2562"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="Serine/threonine specific protein phosphatase"</p> <p>/protein_id="CAF20997.1"</p> <p>/db_xref="GI:41326515"</p> <p>/translation="MYKQIDSLVPTDGESIAKARLASMNTDMWSSDLHLGKRFVAM RGFDPDEDEVILSNFETIGDIDVLMIGDLSGSHRAEERALYLARLGGVYGH LKPNHDSCHPMYRAHYKORRELFVDSVQAKPMKDEDEVILSHPRQDDHGM ESRFPDLRLRVPLLIHGLHLSQPMTPGQVDVGEAWGLKAPRELVLQKLWESLSE KI"</p> <p>830..2722</p> <p>/gene="lcp"</p> <p>/locus_tag="cg2563"</p> <p>830..2722</p> <p>/gene="lcp"</p> <p>/locus_tag="cg2563"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="ECTOINE BETAINE TRANSPORTER"</p> <p>/protein_id="CAF20998.1"</p> <p>/db_xref="GI:41326516"</p> <p>/translation="MSTNSGNMLPESQSPPEEPHYPDTHPGLVPGISVDAQNKGL DKTGVATALLIAPFAMGISSDVSSTSTFSAAMTNTGMLNFMVIGITMLY IAFSEYGRKLGTDDEPEPFRSPTAMFGAGIGVIFPFGSEPLMHVLSPPHYH EBPESRHLQALASHQSHFWGLSAMGVALVGAALVASSVRGRVTLLSFRRLRGK TEGIALRLDMALITATLPGTAATLGLSAIQVQCGVQIISGASEITNNILAIITIT IGFILISVSAGKRIKLYSNLISLTGLVLFVITGPTLFLNLISSVLEGESEFL SNAGKLSWGEETIERQOAWMTAFYAMWIAWTFVEMFTARISGRTRLREFALITVAI PSFIIIAFTIIFCGTAITNNRENVDPDSSSKQVLPDFMSLPLYSITPFIILFVL AVEFTASDASVAVMGWSSQGNPANKLIIVVGGCMGMIIVMMLLTGGSALTGQ NLITIIAIPFALVILYMAIAFIKDLSSTPAALIFORAKAISAIVVAGGEHDDPEL SIERPREGAGATPSTADHITDWIQRIDEBGNDVDYDFTTKMADGWTPESTEGS VDAKD"</p> <p>complement(3014..3295)</p> <p>/locus_tag="cg2564"</p> <p>complement(3014..3295)</p> <p>/locus_tag="cg2564"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="hypothetical protein predicted by</p>	

[illegible]


```

gene      complement(8387. .8911)
          /gene="dctx"
          /locus_tag="cg2569"
          complement(8387. .8911)
          /gene="dctx"
          /locus_tag="cg2569"
          /codon_start=1
          /transl_table=11
          /product="PUTATIVE C4-DICARBOXYLATE TRANSPORT SYSTEM
          (PERMEASE SMALL PROTEIN)"
          /protein_id="CAF21004.1"
          /db_xref="GI:41326522"
          /translacion="MMNFKSIYCVTAMQVFSROVSHSPSTWSEELFWLSPAGS
          AFLFGRGHIAVDPIARKLPVSAORVLQVIVQLLVPAALIGMIWGLVLAASIANQO
          LIALPLITGMVYVVIPIAGVFIALPAIIDLIEVAKKEPYPVLDVSESEPPDLDELA
          OSALDSASSASBERN"
          complement(8908. .9657)
          /gene="dctp"
          /locus_tag="cg2570"
          complement(8908. .9657)
          /gene="dctp"
          /locus_tag="cg2570"
          /codon_start=1
          /transl_table=11
          /product="PUTATIVE C4-DICARBOXYLATE-BINDING PROTEIN"
          /protein_id="CAF21005.1"
          /db_xref="GI:41326523"
          /translacion="MMGSEVERKXYIRSFDMFTTPREVYIGSIHEOMSVIRPDDMEPL
          FQSLAESDNISVYIGFPTQGRNLYTTDDAPVKRPADLAGKRIKQVSAHRIETELMG
          SATPLTVEBNVYIAQSGVLDGAENNEISVYONHEVARVNSNTNHLVGLMYVRHD
          LLDMASEPDRLEFLBEMDPAAMTEHDLMTETEDDAIVERAKGSAEFVDDQAFTDAL
          APIKDFLTSEFQRELYEAVRAADTSGGNAA"
          complement(9722. .11569)
          /gene="1epa"
          /locus_tag="cg2571"
          complement(9722. .11569)
          /gene="1epa"
          /locus_tag="cg2571"
          /codon_start=1

CDS
Query Match      94.8%; Score 5656.2; DB 1; Length 349136;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;

1 AGTCGTGAGCGCCACCATGATGTGATGTCACCGAGCTTGGCGAGCTTCTACATCT 60
344394 AGTCGTGAGCGCCACCATGATGTGATGTCACCGAGCTTGGCGAGCTTCTACATCT 344335
61 ACGCTCCGTCGCGCGTGAAGTGGGGTCAATTACGGGTGGGATCAGCCCGGTGAAAGTTGCG 120
344334 ACGCTCCGTCGCGCGTGAAGTGGGGTCAATTACGGGTGGGATCAGCCCGGTGAAAGTTGCG 344275
121 GAACCAATGCTGCTCTGTGGGTGAGGAAACAGTGGCGGTGAAAGTTTCAAGTGC 180
344274 GAACCAATGCTGCTCTGTGGGTGAGGAAACAGTGGCGGTGAAAGTTTCAAGTGC 344215
181 TCTGCAAGTTTAAAGTTATGATCATCATCACTGGAAGGCTGAGTAATTCAGTAGACCTG 240
344214 TCTGCAAGTTTAAAGTTATGATCATCATCACTGGAAGGCTGAGTAATTCAGTAGACCTG 344155
241 CAACGACGAGCCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAGACATTAAGACATCA 300
344154 CAACGACGAGCCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAGACATTAAGACATCA 344095
301 CGTCTTATGCTTGTCTGGAAGAAACCAATTAACCTCAAGAAAGTGGCAAGTGGTCAAT 360
344094 CGTCTTATGCTTGTCTGGAAGAAACCAATTAACCTCAAGAAAGTGGCAAGTGGTCAAT 344035
361 ATCAAGAAATGAGAGTCAAGAGCTTAAATAAAATTTGAGGGAAGATTTGTTCCCGCCCTCG 420
344034 ATCAAGAAATGAGAGTCAAGAGCTTAAATAAAATTTGAGGGAAGATTTGTTCCCGCCCTCG 343975
421 GGGTATGATGAGCTTTCTCCAACTCGAAGACGGGATCATCAAGAACTCTCTGAGAAC 480

```

```

Db      343974 GGGTATGATGAGCTTTCTCCAACTCGAAGACGGGATCATCAAGAACTCTCTGAGAAC 343915
Qy      481 CAGCACTTAAAAACGAGAGATTCACCCCGAACTCCCAAGATTTTCCCGGTTTATTTG 540
Db      343914 CAGCACTTAAAAACGAGAGATTCACCCCGAACTCCCAAGATTTTCCCGGTTTATTTG 343855
Qy      541 ATCTTCATTAATCAAGGTGGAACCGGTGGCGGTTTCTTACGGGAACGAGGACGAGCGGA 600
Db      343854 ATCTTCATTAATCAAGGTGGAACCGGTGGCGGTTTCTTACGGGAACGAGGACGAGCGGA 343795
Qy      601 GGAACACCGCGCAGATATCACCGCGAATGGACAGACCGGTGATGTTGCAAGCATGTTT 660
Db      343794 GGAATGCCGCGCAGATATCACCGCGAATGGACAGACCGGTGATGTTGCAAGCATGTTT 343735
Qy      661 CGGCGCCGCGTGAAGCATCGGACAGGCGAGGTGGAACACCTTATTCCTTGTGTGAAGAG 720
Db      343734 CGGCGCCGCGTGAAGCATCGGACAGGCGAGGTGGAACACCTTATTCCTTGTGTGAAGAG 343675
Qy      721 TCTGCTGTGCGGCAATTCACCTCGAGGGCCCTTTTCATCAAGCATGCCGTTGTGTCCTC 780
Db      343674 GCTGCTGTGCGGCAATTCACCTCGAGGGCTCTTTTCATCAAGCATGCCGTTGTGTCCTC 343615
Qy      781 AAAACCCGGAATTTCAATTTTCCCGGCAACCAACAGATCTTGCCCGGTTGATCCATGCGG 840
Db      343614 AAAACCCGGAATTTATTTTCCCGGCAACCAACAGATCTTGCCCGGTTGATCCATGCGG 343555
Qy      841 GAAAAGTTGGATCAAAATTCATACAGTACAGGCGCGGAAACATGACAAATCTTTCTGAGCTTC 900
Db      343554 GAAAAGTTGGATCAAAATTCATACAGTACAGGCGCGGAAACATGACAAATCTTTCTGAGCTTC 343495
Qy      901 TCGATCTCTGCGAGCGACACACATCATTTGCTTCTTGCGGCAACATGATCAGATTTTG 960
Db      343494 TCGATCTCTGCGAGCGACACACATCATTTGCTTCTTGCGGCAACATGATCAGATTTTG 343435
Qy      961 ATACCACTACACAGCGCAATTCCTTTGCTTAAAGAAAATGTGACGCTACGCGCTACGC 1020
Db      343434 ATACCACTACACAGCGCAATTCCTTTGCTTAAAGAAAATGTGACGCTACGCGCTACGC 343375
Qy      1021 ATTGTTTAAATGCAATGCTCTCCGCTGCAATCATTAAGGCTCCCGGCGCGTGGCGCTTTGC 1080
Db      343374 ATTGTTTAAATGCAATGCTCTCCGCTGCAATCATTAAGGCTCCCGGCGCGTGGCGCTTTGC 343315
Qy      1081 TTGCTGCGGCAAGTCCGCGGAGACGATATGTTGAGTTGATCGCGACCGCGTGCATTTGG 1140
Db      343314 TTGCTGCGGCAAGTCCGCGGAGACGATATGTTGAGTTGATCGCGACCGCGTGCATTTGG 343255
Qy      1141 CCGATGAAAGGTGATCTAGCTGTTCCAAACACGCTTTTTCATCAACGACGCGCATGCG 1200
Db      343254 CCGATGAAAGGTGATCTAGCTGTTCCAAACACGCTTTTTCATCAACGACGCGCATGCG 343195
Qy      1201 AAGCGCGGGAATGCAACGATGATCATTTTGGGGGTTTGAAGTCAACCGACCG 1260
Db      343194 AAGCGCGGGAATGCAACGATGATCATTTTGGGGGTTTGAAGTCAACCGACCG 343135
Qy      1261 ATGAGATGCGCCGCTGCGCGATGCGCGGCATCGCGGCGGCAACGACATGAGGA 1320
Db      343134 ATGAGATGCGCCGCTGCGCGATGCGCGGCATCGCGGCGGCAACGACATGAGGA 343075
Qy      1321 GTCAAGTGTGACACAGTGGCGAGGGTATGACGCTTATGACGCGACCTCCACACT 1380
Db      343074 GTCAAGTGTGACACAGTGGCGAGGGTATGACGCTTATGACGCGACCTCCACACT 343015
Qy      1381 CAACCGTCCGCGCTAAATTTCTCGCTTGTGCGATCAAGAAATGCTAAATCCAAACCTG 1440
Db      343014 CAACCGTCCGCGCTAAATTTCTCGCTTGTGCGATCAAGAAATGCTAAATCCAAACCTG 342955
Qy      1441 CAATTTTGTGATCTTGAATCAACGCGCAGGTGCAAAAGTCCATTTAGTCAATCAAG 1500
Db      342954 TAAATTTTGTGATCTTGAATCAACGCGCAGGTGCAAAAGTCCATTTAGTCAATCAAG 342895
Qy      1501 TACTTAAATGATCAAGTAAACTATCTGAT-TTTAAAGAGTCCCAACATGAAATCACT 1559

```


Db 342894 TAATTTAAATACGAGCAAACTTTCCTGATTAATAAAGAGTCCGACATGACATCATC 342835
Qy 1560 ATCTGAAAGACGAGCAAGATCGGCAAGCGATTGCACTCTTAATCGCACTCTTCGCC 1619
Db 342834 ATCTGAAAGACGAGCAAGATCGGCAAGCGCGCAGCCCTGATCGCACTCTTCGCA 342775
Qy 1620 AACCAAGGATGAACTTTGGGGCTTGGCAACAGATCTCTACCACTGAGTACTCAAGAG 1679
Db 342774 ACTAAGGGGGGAACCTTTGGGGCTTGGCACTGAGTCTCTACCTTTGAGCACTCAAGAG 342715
Qy 1680 CTCATTCGATGTATGAAAGCTGGGGAAGTGTCAATTCGAAGTCAAGGCAATTCCTGTTG 1739
Db 342714 CTCATTCGATGTATGAAAGCTGGGGAAGTGTCAATTCGAAGTCAAGGCAATTCCTGTTG 342655
Qy 1740 GATGAATTAAGTGGGACTTAACCCGTGACGATGAAAAAGCTACTTTTAAACATTCGCCAA 1799
Db 342654 GATGAATTAAGTGGGACTTAACCCGTGACGATGAAAAAGCTACTTTTAAACATTCGCCAA 342595
Qy 1800 GAGTCACTGACCAATCGACATCGTTGATGAAGAGTCTACAGCCGAGATGGTGGCAAC 1859
Db 342594 GAGTCACTGACCAATCGACATCGTTGATGAAGAGTCTACAGCCGAGATGGTGGCAAC 342535
Qy 1860 CCTGATCCATACGAGACAGCTGCAAGTATGAGGCAAAAGATCGCTGCAAGATCCGTTGAA 1919
Db 342534 CCTGATCCATACGAGACAGCTGCAAGTATGAGGCAAAAGATCGCTGCAAGATCCGTTGAT 342475
Qy 1920 GTTCAAAATCCTTGGGATCGGCGGAAACCG - CACATGCTTTCAATTGAACATATCTTCT 1978
Db 342474 GTTCAAAATCCTTGGGATCGGCGGAAACCG - CACATGCTTTCAATTGAACATATCTTCT 342415
Qy 1979 CTGTGAGACTGACAAAGGCTCGAGGCGTGCACCTTAAACGTGGAGGCAACGCTGCA 2038
Db 342414 CTGTGAGACTGACAAAGGCTCGAGGCGTGCACCTTAAACGTGGAGGCAACGCTGCA 342355
Qy 2039 TTCTTCAACACCATCGAAGAGTCCCAACCAAGCCGTGCAACCCAGGGTTTGGGCACTTTG 2098
Db 342354 TTCTTCAACACCATCGAAGAGTCCCAACCAAGCCGTGCAACCCAGGGTTTGGGCACTTTG 342295
Qy 2099 TCCCGGCGCCAAACATCGTGTGTGTGGCACTGTGGAAAGAAAGCCGACCGCATCCGC 2158
Db 342294 TCCCGGCGCCAAACATCGTGTGTGTGGCACTGTGGAAAGAAAGCCGACCGCATCCGC 342235
Qy 2159 GGAACGTGGAAGGCCAGTGAATGCTTCTTGGCCAGGTTCCATCTCTGATGATGACAAAC 2218
Db 342234 GGAACGTGGAAGGCCAGTGAATGCTTCTTGGCCAGGTTCCATCTCTGATGATGACAAAC 342175
Qy 2219 ATGCCACATCATCTGTGATGAGCAGAGATATCCAACTGGAAAAAGCTGATCACTAC 2278
Db 342174 ATGCCACATCATCTGTGATGAGCAGAGATATCCAACTGGAAAAAGCTGATCACTAC 342115
Qy 2279 CGTTCATGAGCAATTAAGTGTGCGCTTGAAACAAAGAAAGTACTGTGTGGGCTTA 2338
Db 342114 CGTTCATGAGCAATTAAGTGTGCGCTTGAAACAAAGAAAGTACTGTGTGGGCTTA 342055
Qy 2339 TGCACACAGAACTTTCCAGTTTGCCTGCGCTGCGATGATGATCTCTCCGACGGGCGAGG 2398
Db 342054 TGCACACAGAACTTTCCAGTTTGCCTGCGCTGCGATGATGATCTCTCCGACGGGCGAGG 341995
Qy 2399 TCAATGATCCAAAGGATGTAAGTGTGATGAGTAACTCTCCAGCTCTACTACAGACAG 2458
Db 341994 TCAATGATCCAAAGGATGTAAGTGTGATGAGTAACTCTCCAGCTCTACTACAGACAG 341935
Qy 2459 ATCCAGGTTTCCCTTGCACCAAGCGCAACGGCTGGGCTGACACCAACGCGCTTGA 2518
Db 341934 ATCCAGGTTTCCCTTGCACCAAGCGCAACGGCTGGGCTGACACCAACGCGCTTGA 341875
Qy 2519 CCGGACCGGACGATTTGAGTGGACGACCTGCGGACGCTCTTTACCCGGATGATCTCT 2578
Db 341874 CCGGACCGGACGATTTGAGTGGACGACCTGCGGACGCTCTTTACCCGGATGATCTCT 341815
Qy 2579 ATGACCTGATGATGATCTATTCGATGAGCGGATATTTACTGACGACACTTAAACTTT 2638
Db 341814 ATGACCTGATGATGATCTATTCGATGAGCGGATATTTACTGACGACACTTAAACTTT 341755

Qy 2639 TCTACACCGGCAACCTTAAATTTGACGGAAGCGCGCCGACCCCAAACTTGTGGAG 2698
Db 341754 TCTACACCGGCAACCTTAAATTTGACGGAAGCGCGCCGACCCCAAACTTGTGGAG 341695
Qy 2699 TCGAGAACCCAACTGGGCTGATGAGCGGCAATTCATGCGCGTTGCGCTTAAATTCGCTTA 2758
Db 341694 TCGAGAACCCAACTGGGCTGATGAGCGGCAATTCATGCGCGTTGCGCTTAAATTCGCTTA 341635
Qy 2759 TCGAGAACCCGCGCAGCGGTTTCAACCCCATTTACCGGATCCATGATGAGCCTGATG 2818
Db 341634 TCGAGAACCCGCGCAGCGGTTTCAACCCCATTTACCGGATCCATGATGAGCCTGATG 341575
Qy 2819 GTGATGTTGGAACATGTTCTTTGGGGCCCAACGCGAAACCTCACCCGTTGACGCGTTTC 2878
Db 341574 GTGATGTTGGAACATGTTCTTTGGGGCCCAACGCGAAACCTCACCCGTTGACGCGTTTC 341515
Qy 2879 TATACCGCTGACAGATCTTGAAGAACTGGGAATTCCTCGGTGAATCACTTTGACCTCA 2938
Db 341514 TATACCGCTGACAGATCTTGAAGAACTGGGAATTCCTCGGTGAATCACTTTGACCTCA 341455
Qy 2939 GTGATGACAACTGGTTCTGCTCTGATCTGCTTCCGATGAGCTACATGTGGGAATGCC 2998
Db 341454 GTGATGACAACTGGTTCTGCTCTGATCTGCTTCCGATGAGCTACATGTGGGAATGCC 341395
Qy 2999 CCAACCTTTTACGCTTCCGCAATGAAAGAACTGGGCAAAATCTCGACGTTGATTTTCT 3058
Db 341394 CCAACCTTTTACGCTTCCGCAATGAAAGAACTGGGCAAAATCTCGACGTTGATTTTCT 341335
Qy 3059 GTCCAAAGATTTGACCGAATCCAGATGAGGTTTACTACTACGCAACTGACCACT 3118
Db 341334 GTCCAAAGATTTGACCGAATCCAGATGAGGTTTACTACTACGCAACTGACCACT 341275
Qy 3119 GCGGATATGTCGTGCAAAAGCTTGAAGAAAGCACTTCCGCTTTGCGAGATTCAGCG 3178
Db 341274 GCGGATATGTCGTGCAAAAGCTTGAAGAAAGCACTTCCGCTTTGCGAGATTCAGCG 341215
Qy 3179 AGCTGATTTCCGCGCATGAAATCTACAGCACCGCAAGTTGCACTTGAATGCT 3238
Db 341214 AGCTGATTTCCGCGCATGAAATCTACAGCACCGCAAGTTGCACTTGAATGCT 341155
Qy 3239 GCGTCCGCGGCTGATGAGGCGCTGCCGCGCAGAGATGATCAACCAAGTTGACAGAGAG 3298
Db 341154 GCGTCCGCGGCTGATGAGGCGCTGCCGCGCAGAGATGATCAACCAAGTTGACAGAGAG 341095
Qy 3299 GATGGGTGACATGCTGATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358
Db 341094 GATGGGTGACATGCTGATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 341035
Qy 3359 AAGAGCTCTTCTCCCAAGGGGAGTGGGGGATATCAAGTCTGATTAAGTGTGCAAC 3418
Db 341034 AAGAGCTCTTCTCCCAAGGGGAGTGGGGGATATCAAGTCTGATTAAGTGTGCAAC 340975
Qy 3419 CTGTCCGATGACATCCGAGGCAATATTTCCCTCGAGTGGGATGATGTCGCTTGTCTG 3478
Db 340974 CTGTCCGATGACATCCGAGGCAATATTTCCCTCGAGTGGGATGATGTCGCTTGTCTG 340915
Qy 3479 TCGATCGTATGATGATGTCGCTGCGATGAGTGAATCTTGGCAATTAAGTATGCGCG 3538
Db 340914 TCGATCGTATGATGATGTCGCTGCGATGAGTGAATCTTGGCAATTAAGTATGCGCG 340855
Qy 3539 ACGATTAATACACCATTTGATGATTAACCTGAGTGTGAGACAGGTTTCACTTTCGCG 3598
Db 340854 ACGATTAATACACCATTTGATGATTAACCTGAGTGTGAGACAGGTTTCACTTTCGCG 340796
Qy 3599 GCGTTCAAAGGTGACATTTTGAAGATTAAGTATTAAGGATCTTTTGTGGCAATTT 3658
Db 340795 GCGTTCAAAGGTGACATTTTGAAGATTAAGTATTAAGGATCTTTTGTGGCAATTT 340736
Qy 3659 GTACAAATATCTTGCAGAAATCCCTTGTATGACACAAATTAACAGGTTTAATATTTGTTTA 3718
Db 340735 GTACAAATATCTTGCAGAAATCCCTTGTATGACACAAATTAACAGGTTTAATATTTGTTTA 340676

QY 3719 GCTTTGAAACAATCATGCTGAATATTTTGTCTTCCGGTTAAGGAAATTC 3778
 DB 34075 GCTTTGAAACAATCATGCTGAATATTTTGTCTTCCGGTTAAGGAAATTC 340616
 QY 3779 ATGACCAATTAAGACCTCGCGCAAGCAATCCGCGCAATTTGGGGGAGAGACAAT 3838
 DB 340615 ATGACCAATTAAGACCTCGCGCAAGCAATCCGCGCAATTTGGGGGAGAGACAAT 340556
 QY 3839 GTCGCGCGCGCACACTGTGCAACGCGTTTACGCTGTGCTCAAGACAACCAAGATGTG 3898
 DB 340555 GTCGCGCGCGCACACTGTGCAACGCGTTTACGCTGTGCTCAAGACAACCAAGATGTG 340496
 QY 3899 GATCGCAAAAGTCTGATGATGATTCAGATCTGAAAGGCACTTTGAAACTGGCGGCAATG 3958
 DB 340495 GATCGCAAAAGTCTGATGATGATTCAGATCTGAAAGGCACTTTGAAACTGGCGGCAATG 340436
 QY 3959 TTCGAGATCATGTCGGGCGCAGGCGATGTGATCATGTTTCAAGAACTCGATGACGA 4018
 DB 340435 TTCGAGATCATGTCGGGCGCAGGCGATGTGATCATGTTTCAAGAACTCGATGACGA 340376
 QY 4019 ACCTCGAAAGACATCGCTGTGTCACAGAGACGCTCAAGATGTTGTGCTAACAACGCC 4078
 DB 340375 ACCTCGAAAGACATCGCTGTGTCACAGAGACGCTCAAGATGTTGTGCTAACAACGCC 340316
 QY 4079 AACTGCTTCAAGCCGTGCTGTGAAGGATTTGGGGGCAATTTGTCCCGCTGATCCATTC 4138
 DB 340315 AACTGCTTCAAGCCGTGCTGTGAAGGATTTGGGGGCAATTTGTCCCGCTGATCCATTC 340256
 QY 4139 TTGGTTGTGGCGGCTGTGCTCATGAGCTATCAACATGTGTGTTGGCGAGATCTGTTC 4198
 DB 340255 TTGGTTGTGGCGGCTGTGCTCATGAGCTATCAACATGTGTGTTGGCGAGATCTGTTC 340196
 QY 4199 GGTCCGCAATCACTGCTGTGAAGATTTCCCTCAGATGAGCGGTGTGCTGAATGATCAAC 4258
 DB 340195 GGTCCGCAATCACTGCTGTGAAGATTTCCCTCAGATGAGCGGTGTGCTGAATGATCAAC 340136
 QY 4259 CTGATGGCAATGCGCGCGTTGCGGCTTGGCCAGTGTGTTGTTGCTTCCGCAACCAAG 4318
 DB 340135 CTGATGGCAATGCGCGCGTTGCGGCTTGGCCAGTGTGTTGTTGCTTCCGCAACCAAG 340076
 QY 4319 CGTTTCGTGGCAATGATGATTTCTTGCGCGCGCGCATTTGTATGGCGATGTGTTCCCAAC 4378
 DB 340075 CGTTTCGTGGCGAATGATGATTTCTTGCGCGCGCGCATTTGTATGGCGATGTGTTCCCAAC 340016
 QY 4379 CTGCTTAAAGGCTACGACTGTGCGCGCGCAACCATACCGCGCGGCAATGTCATATGTGTC 4438
 DB 340015 CTGCTTAAAGGCTACGACTGTGCGCGCGCGCAACCATATGTCATATGTGTC 339956
 QY 4439 CTGTTTGGTTTGAATGTTGCTCAAGCTGTTACAGAGGCAACGCTTCTGTGCTGTG 4498
 DB 339955 CTGTTTGGTTTGAATGTTGCTCAAGCTGTTACAGAGGCAACGCTTCTGTGCTGTG 339896
 QY 4499 GTCTCTGATTTCTGCAACGATCGAAGATTTCTGCAACAGGCACTCATGAGCACTGCA 4558
 DB 339895 GTCTCTGATTTCTGCAACGATCGAAGATTTCTGCAACAGGCACTCATGAGCACTGCA 339836
 QY 4559 GACTTCTGATCAACCCAGTGTGACTCTGCTGCTGCAACCGGCTTCTTACGTTCAATTTGCT 4618
 DB 339835 GACTTCTGATCAACCCAGTGTGACTCTGCTGCTGCAACCGGATTTCTTACATTCATGCTGC 339776
 QY 4619 ATTGTCGAGCAATGCGTGGTGGTGAATTTGCTGCAACAGGCTGCAAGGAGACTTAT 4678
 DB 339775 ATTGTCGAGCAATGCGTGGTGGTGAATTTGCTGCAACAGGCTGCAAGGAGACTTAT 339716
 QY 4679 GATTTGCGTGGTTCAGTGGCGGTCTGCTTTTGTGCTGTGCTTACTCAACCAATGCTTATC 4738
 DB 339715 GATTTGCGTGGTTCAGTGGCGGTCTGCTTTTGTGCTGTGCTTACTCAACCAATGCTTATC 339656
 QY 4739 ACTGCTTGAACAGAGCTTCCCGCAATGAGCTGGAGCTGTCAACAGGGTGAATCC 4798
 DB 339655 ACTGCTTGAACAGAGCTTCCCGCAATGAGCTGGAGCTGTITTAACAGGGTGAATCC 339596
 QY 4799 TTCACTTTCGAAACCGCATCATGCGCAATATGCGCGAGGTTGCAAGATTTTGGCAGTG 4858

DB 339595 TTCACTTTCGAAACCGCATCATGCTAATATATGCGCCAGGGTGGGATGTTTGGCAGTG 339536
 QY 4859 TTCTTCTTCAGGAAGATGAAAGCTCAAGGGCTTCAAGTGTCTTCAAGTGTCTCCGCT 4918
 DB 339535 TTCTTCTTCAGGAAGATGAAAGCTCAAGGGCTTCAAGTGTCTTCAAGTGTCTCCGCT 339476
 QY 4919 GTTCTTGGTATTAAGAGACCTGCGAATCTTGGTGTGAACCTTGGCTGGCGCTGCGCTTC 4978
 DB 339475 GTTCTTGGTATTAAGAGACCTGCGAATCTTGGTGTGAACCTTGGCTGGCGCTGCGCTTC 339416
 QY 4979 TACATTTGGTATCGGTATCGCAGCTATGCTGCGCTTGTGATTCACCTTTGATATCAAG 5038
 DB 339415 TACATTTGGTATCGGTATCGCAGCTATGCTGCGCTTGTGATTCACCTTTGATATCAAG 339356
 QY 5039 GCAGTTGCGTTGGCGCGCTGCAAGTTTCTTGGGTGTGTTTCTATGATATGCTCCAGATATG 5098
 DB 339355 GCAGTTGCGTTGGCGCGCTGCAAGTTTCTTGGGTGTGTTTCTATGATATGCTCCAGATATG 339296
 QY 5099 GTCAATGTTCTGTTTGGCGGATGTTAATCTTGTATCGCATTCGGCGCAGCATTCCT 5158
 DB 339295 GTCAATGTTCTGTTTGGCGGATGTTAATCTTGTATCGCATTCGGCGCAGCATTCCT 339236
 QY 5159 TATGCGCTTTACTTGTGTCGCGCAACGCGCAATGATCCAGATGCAACCGCTGCTCA 5218
 DB 339235 TATGCGCTTTACTTGTGTCGCGCAACGCGCAATGATCCAGATGCAACCGCTGCTCA 339176
 QY 5219 GTGCTGCAAGAGACCAACAAAGCCGGAAGACAACCCGAGAAATTTCAAGATTC 5278
 DB 339175 GTGCTGCAAGAGACCAACAAAGCCGGAAGACAACCCGAGAAATTTCAAGATTC 339116
 QY 5279 ACCATCATCCAGGCACTTTGACCGGTGAACCTATGCACTGACAGCGTCAGCGATGCC 5338
 DB 339115 ACCATCATCCAGGCACTTTGACCGGTGAACCTATGCACTGACAGCGTCAGCGATGCC 339056
 QY 5339 ATGTTTGCACGCGAAGACTTGGCTCAGGTGTGCGATGTTCCCAACCAAGGGCAGTG 5398
 DB 339055 ATGTTTGCACGCGAAGACTTGGCTCAGGTGTGCGATGTTCCCAACCAAGGGCAGTG 338996
 QY 5399 GTTTCACAGAGAGCGGAAAGATGCTGTGCGCTTCCCATCTGATCGCTTGGCAGTC 5458
 DB 338995 GTTTCACAGAGAGCGGAAAGATGCTGTGCGCTTCCCATCTGATCGCTTGGCAGTC 338936
 QY 5459 CGCACTAAGGCTGAGAGATGTTCCATGATGATATCTTGAATGCAATTTGTTGCAACC 5518
 DB 338935 CGCACTAAGGCTGAGAGATGTTCCATGATGATATCTTGAATGCAATTTGTTGCAACA 338876
 QY 5519 GTTAACTTCAACGCGACGCACTTTAACCGCTGAAAGACAGGGGATGAAATCAAGCA 5578
 DB 338875 GTTAACTTCAACGCGACGCACTTTAACCGCTGAAAGACAGGGGATGAAATCAAGCA 338816
 QY 5579 GGGGAGCTGCTGTGGAATTTGATATGATGCAATTAAGGCTGCAAGTTATGAGGTAAC 5638
 DB 338815 GGGGAGCTGCTGTGGAATTTGATATGATGCAATTAAGGCTGCAAGTTATGAGGTAAC 338756
 QY 5639 ACGCGATTTGTTTTCGAATTAACAAGAAACCGGACCTGTAAACACTTACGTTTGGGC 5698
 DB 338755 ACGCGATTTGTTTTCGAATTAACAAGAAACCGGACCTGTAAACACTTACGTTTGGGC 338696
 QY 5699 GAAATTTGAAGCGGAGCCAACTGCTCAACGTCGAAAGAAAGAAAGCGGTGCACAGCA 5758
 DB 338695 GAAATTTGAAGCGGAGCCAACTGCTCAACGTCGAAAGAAAGAAAGCGGTGCACAGCA 338636
 QY 5759 CCATAAGTTGAACCTTGAAGTTTGCACAACAGTTTGAAGTGAAGTGAAGTGAAGTGAAG 5818
 DB 338635 CCATAAGTTGAACCTTGAAGTTTGCACAACAGTTTGAAGTGAAGTGAAGTGAAGTGAAG 338576
 QY 5819 ATCTTTGACACCGGATACCGGATACGCTTCAAGATTTTAAACCTGTTCAACAGGTCATGC 5878
 DB 338575 ATCTTTGACACCGGATACCGGATACGCTTCAAGATTTTAAACCTGTTCAACAGGTCATGC 338516
 QY 5879 TCGGTGTACTGTGTGTGTCACCCCGCAATTTTCAACCCCATTTGGAATGTTGTTCA 5938

Db 338515 TCGGTGTACCTGTGTGTGCGACCCCGCAATCTTCAACCCACATGTGATCTGTCA 338456
QY 5939 GCAAGTACGTTTATATATTTTGGCCCGCTG 5967
Db 338455 GCAGTAGCGTTGTATATTTTGGCCCGCTG 338427

RESULT 4
AX127152/c 349980 bp DNA linear PAT 11-MAY-2001
LOCUS AX127152 Sequence 7068 from Patent EP108790.
DEFINITION AX127152 AX127152
ACCESSION AX127152.1 GI:14041140
VERSION AX127152.1
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,
TITLE Novel polynucleotides
JOURNAL Patent: EP 108790-A 7068 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source
1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:11718"
/note="Seq 1 to long (3,309,400) split in 11, seq 7068
2,700,001 3,049,980"

ORIGIN
Query Match 94.8%; Score 5656.2; DB 6; Length 349980;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;
QY 1 AGTCCGTGACGCGCATTGATGTGTGTGTCACGAGCTTGCGAGGCTTTCTAATCT 60
Db 115585 AGTCCGTGACGCGCATTGATGTGTGTGTCACGAGCTTGCGAGGCTTTCTAATCT 115526
QY 61 ACGCTCCCGCTGCGCGTGTGAGTGGGTCAATTCACGCGTGGATCAACGCGGTGAAGTTGCG 120
Db 115525 ACACGTGCGTGGCGGTGTGAGTGGGTCAATTCACGCGTGGATCAACGCGGTGAAGTTGCG 115466
QY 121 GAACCATGTGTTCTTGTGTGTGAGGAAAGAGTGTGCGGTGAGAGTTTTCAGATG 180
Db 115465 GAACCATGTGTTCTTGTGTGTGAGGAAAGAGTGTGCGGTGAGAGTTTTCAGATG 115406
QY 181 TCTGCAAGTTTAAAGTATGATCATCATGCTTGGAAAGGCTGAGATTAATCAAGACCTG 240
Db 115405 TCTGCAAGTTTAAAGTATGATCATCATGCTTGGAAAGGCTGAGATTAATCAAGACCTG 115346
QY 241 CAACAGCAGGCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAGACATAGACATCAT 300
Db 115345 CAACAGCAGGCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAGACATAGACATCAT 115286
QY 301 CGTCTATGTCTTGTGTGAAAGAAACCAATTAACCTCAAGAAAGTGCAGAAAGTGTGCATT 360
Db 115285 CGTCTATGTCTTGTGTGAAAGAAACCAATTAACCTCAAGAAAGTGCAGAAAGTGTGCATT 115226
QY 361 ATCAAGAAATGCAAGTCAAGCAATTAAGAAATGAAGGAAGATTTTCCCGCCCTCG 420
Db 115225 ATCAAGAAATGCAAGTCAAGCAATTAAGAAATGAAGGAAGATTTTCCCGCCCTCG 115166
QY 421 GGGTATGATGCTTTCTCAACTCGAAACGCGCATCATCAAGAACTCTCTGAGAAC 480
Db 115165 GGGTATGATGCTTTCTCAACTCGAAACGCGCATCATCAAGAACTCTCTGAGAAC 115106
QY 481 CAGCACTTAAAAACGAGATTCACCCCGAACTCCCGCAGATTGTTCCCGGTTTTATTG 540
Db 115105 CAGCACTTAAAAACGAGATTCACCCCGAACTCCCGCAGATTGTTCCCGGTTTTATTG 115046

QY 541 ATCTTCAATATGACGGTGGAAACGGTGGCGGCTTTCTTAAGGGAACCGAGGACCAAGGCGA 600
Db 115045 ATCTTCAATATGACGGTGGAAACGGTGGCGGCTTTCTTAAGGGAACCGAGGACCAAGGCGA 114986
QY 601 GGAACACCGCGAGTATCACCGCGAACAATGACGACGCGTGTATGTTCCCAACATGTTT 660
Db 114985 GGAATGCGCGAGTATCACCGCGAACAATGACGACGCGTGTATGTTCCCAACATGTTT 114926
QY 661 CGGCGCGCGCTACCGCATGCGGACGCGAGTGGAAACCTTATTTCCCTTGTGTGAAGAG 720
Db 114925 CGGCGCGCGCTACCGCATGCGGACGCGAGTGGAAACCTTATTTCCCTTGTGTGAAGAG 114866
QY 721 TCGTGTGTGGGCAATTCACCTGAGGGGCTTTATCAAGGCAATGCGGTGGGCTTC 780
Db 114865 GCGTGTGTGGGCAATTCACCTGAGGGGCTTTATCAAGGCAATGCGGTGGGCTTC 114806
QY 781 AAAACCGGATTTATTTTTCCCGGCAACCAACAGATCTTCCGCGGTATCATGCGG 840
Db 114805 AAAACCGGATTTATTTTTCCCGGCAACCAACAGATCTTCCGCGGTATCATGCGG 114746
QY 841 GAAAAGTTGATCAATGATGATCAAGTACGCGCGGAACTGACATCTTTGTAGCTTC 900
Db 114745 GAAAAGTTGATCAATGATGATCAAGTACGCGCGGAACTGACATCTTTGTAGCTTC 114686
QY 901 TCGATCTGCGCAGCGACCAATCATTTGCTTCTTGGGCACTGATGAGATTTTG 960
Db 114685 TCGATCTGCGCAGCGACCAATCATTTGCTTCTTGGGCACTGATGAGATTTTG 114626
QY 961 ATACCACTACACGCGCAATTTGCTTGAAGAAATGTAAGCGGTACGCGTACGCG 1020
Db 114625 ATACCACTACACGCGCAATTTGCTTGAAGAAATGTAAGCGGTACGCGTACGCG 114566
QY 1021 ATTTGTTCAATGCAATGCTCCGCTGATCATAGGCGTCCCGGACGCTGGCGCTTTC 1080
Db 114565 ATTTGTTCAATGCAATGCTCCGCTGATCATAGGCGTCCCGGACGCTGGCGCTTTC 114506
QY 1081 TTGCTGCGCAGTGTGCGGCGGAGCGATATGTTGATGATGCGGACGCGGTGATTTGG 1140
Db 114505 TTGCTGCGCAGTGTGCGGCGGAGCGATATGTTGATGATGCGGACGCGGTGATTTGG 114446
QY 1141 CCGATGGAACGCTGATAGTCTGTTCCCAACGCGCTTTTCATCAACGACGCGCATG 1200
Db 114445 CCGATGGAACGCTGATAGTCTGTTCCCAACGCGCTTTTCATCAACGACGCGCATG 114386
QY 1201 AAGCGCGGAAATGCGAAGCGGTAGTACATTTGGCGTTTGAACGTCAACGCTGACCG 1260
Db 114385 AAGCGCGGAAATGCGAAGCGGTAGTACATTTGGCGTTTGAACGTCAACGCTGACCG 114326
QY 1261 ATGAGTGGCGCGTGTGGCGATGCGGCGGCGCATGCGCGGCGGACCAAGCACTAGCGA 1320
Db 114325 ATGAGTGGCGCGTGTGGCGATGCGGCGGCGCATGCGCGGCGGACCAAGCACTAGCGA 114266
QY 1321 GTCAAGTGTGACACGACGTGCGCAGGGGTATGACGCTTATCGACGCGACCTCAACCT 1380
Db 114265 GTCAAGTGTGACACGACGTGCGCAGGGGTATGACGCTTATCGACGCGACCTCAACCT 114206
QY 1381 CAACGCGCGCGCTAAATTTCTCGGTCTTGGCGATCAAGAAATGCTTAATCAACCTTG 1440
Db 114205 CAACGCGCGCGCTAAATTTCTCGGTCTTGGCGATCAAGAAATGCTTAATCAACCTTG 114146
QY 1441 CAATTTTGTGTCTTGAATCAACGCGCAGGTGCAAAAGGTCATTAGGTCAACG 1500
Db 114145 CAATTTTGTGTCTTGAATCAACGCGCAGGTGCAAAAGGTCATTAGGTCAACG 114086
QY 1501 TACTTAAATGCAAGTAAATCTATCTGAT-TTTAAAGAGTCCACCATGGAATCACT 1559
Db 114085 TACTTAAATGCAAGTAAATCTTCTGATATATAAGAGTCCGACCATGGAATCACTATC 114026
QY 1560 ATCTGCAAAAGCAGCAAGAGTGGCGAAAGCTTGAAGTCTTAATGCAACCTTTCGCC 1619
Db 114025 ATCTGCAAAAGCAGCAAGAGTGGCGAAAGCTTGAAGTCTTAATGCAACCTTTCGCC 113966
QY 1620 AACAGGGTGAACCTTGGGGCTTGCAACAGGATCCTCACCACTGATACCTACCAAGG 1679

Db	113965	ACTAAGGGGGGAACCTGGGGGCTTGAACTGGATGCTCACTTTGAGCACTTACCAAGG	113906
QY	1680	CTCATTCGCATGTATGAAGCTGGGGAGAGTGCATTTCAAGAACTGCAGAGCATTTCTGTTG	1739
Db	113905	CTCATTCGCATGTATGAAGCTGGGGAGAGTGCATTTCAAGAACTGCAGAGCATTTCTGTTG	113846
QY	1740	GATGAATAACGGGGACTAACCCGCTGACCAATGAAAACAGTACTTTTAAACCAATTCGCAAA	1799
Db	113845	GATGAATAACGGGGAGTTAAACGGCGACAGATGAAAACAGTACTTTCAAAACCAATTCGTA	113786
QY	1800	GAGTTCACTGACACATCGACATCTGTTGATGAAAGGTCAGTCAAGCCACAGATGGTGCMAAC	1859
Db	113785	GAGTTCACTGACACATCGACATCTGTTGATGAAAGGTCAGTCAAGCCACAGATGGTGCMAAC	113726
QY	1860	CCTGATTCATACGAAGCAGCTGCAGAGTATGAGGCAAAAGATGCTGCAGAAATCCGTTGAA	1919
Db	113725	CCTGATTCATACGAAGCAGCTGCAGAGTATGAGGCAAAAGATGCTGCAGAAATCCGTTGAT	113666
QY	1920	GTTCAAAATCCTTGGGACATGGCGGAAACCGG-CACATGCTTTCAATTGAAACATCATCTTCT	1978
Db	113665	GTTCAAAATCCTTGGGACATGGCGGAAACCGGCAATGCTTTCAATGAGCATCATCTTCT	113606
QY	1979	CTGTACAGACGTACAAAGGTCAGGCGCTGCACCTTAAACTGTGAGGACAAACGCTGCA	2038
Db	113605	CTGTACAGACGTACAAAGGTCAGGCGCTGCACCTTAAACTGTGAGGACAAACGCTGCA	113546
QY	2039	TTCCTTCAACACATGAGAAAGGTCCTCAACCGCGCTACCCAGGCTTTGGGCACTTTG	2098
Db	113545	TTCCTTCAACACATGAGAAAGGTCCTCAACCGCGCTACCCAGGCTTTGGGCACTTTG	113486
QY	2099	TCCCGCGCGCAAAACATCGTGTGGTGAGCACTGTGAAAGAAAGCCGACATCCGC	2158
Db	113485	TCCCGCGCGCAAAACATCGTGTGGTGAGCACTGTGAAAGAAAGCCGACATCCGC	113426
QY	2159	GGAACCTGTGAGAGGCCCACTGACTGTTCTTGGCCGAGTTTCCATCCTGTATGATGCAAC	2218
Db	113425	GGAACCTGTGAGAGGCCCACTGACTGTTCTTGGCCGAGTTTCCATCCTGTATGATGCAAC	113366
QY	2219	ATGCACACATCATCGTGTGATGAGAGCAGCAGTATCCAACTGGAAAACGCTGATCATAC	2278
Db	113365	ATGCACACATCATCGTGTGATGAGAGCAGCAGTATCCAACTGGAAAACGCTGATCATAC	113306
QY	2279	CGTCTCATGAGCAATTTAAAGCTGCGCTTGAAACAAAAGAAAGTATGTGTGGGGCTA	2338
Db	113305	CGTCTCATGAGCAATTTAAAGCTGCGCTTGAAACAAAAGAAAGTATGTGTGGGGCTA	113246
QY	2339	TGCACACAGAACTTTCCAGTTTGGCGGCTGACGATGTAATCTCTCCGACAGGCGAGGC	2398
Db	113245	TGCACACAGAACTTTCCAGTTTGGCGGCTGACGATGTAATCTCTCCGACAGGCGAGGC	113186
QY	2399	TCAATGATCCCAACGGAATGTATCGTATGAGATATCCTCAAGTCTACTACAGACG	2458
Db	113185	TCAATGATCCCAACGGAATGTATCGTATGAGATATCCTCAAGTCTACTACAGACG	113126
QY	2459	ATCCAGGTTTCCCTTGCAGCAAAAGCGACCGGCGGCTCAACACACAGCGGCTGA	2518
Db	113125	ATCCAGGTTTCCCTTGCAGCAAAAGCGACCGGCGGCTCAACACACAGCGGCTGA	113066
QY	2519	CCGGAACCGAGGATTTGCAGTGGACACACTGCGCGACGCTCTTTACCCGAGTACATCT	2578
Db	113065	CCGGAACCGAGGATTTGCAGTGGACACACTGCGCGACGCTCTTTACCCGAGTACATCT	113006
QY	2579	ATGACCTGGATGATGCTATTCGGGTGTGAGCCGTAATTTACTGACGGCACCTTAAACTTT	2638
Db	113005	ATGACCTGGATGATGCTATTCGGGTGTGAGCCGTAATTTACTGACGGCACCTTAAACTTT	112946
QY	2639	TCTACACCGGCAACTTAAAAATTTAGACGAAAGCGCGCGCACCCCAAAACCTTGTGAAAG	2698
Db	112945	TCTACACCGGCAACTTAAAAATTTAGACGAGCGCGCGCACCCCAAAACCTTGTGAAAG	112886
QY	2699	TGAGAGACCCAACTGGGCTGATGAGGCGGCAATTCATGCGCGTTGCTTAAAAATTCGCTTA	2758

Db	112885	TCGAGGACCCAACTGGGGCTGATGGCGGGGATTCATCGCTGGCTCTAAAATTCGGCTTA	1128826
Qy	2759	TCGACGACACCGCCAGCGGTTTCAACCCCATTAACCGGAGATCCCATGATCAACCCCTGAAG	2818
Db	112825	TCGAGGAGACCGCCAGCGGTTTCAACCCCATTAACCGGAGATCCCATGATCAACCCCTGAAG	112766
Qy	2819	GTGATGATTGGAAACATGATTCTTTGGGGGCCAACGCGAAAACCTCAACCGGTGACGGGTTTC	2878
Db	112765	GTGATGATTGGAAATGGTTCTTTGGGGGCCAACGCGAAAACCTCAACCGGTGACGGGTTTC	112706
Qy	2879	TATACCGCTCGACAGATCTTGAATACTGGGAATTCCTCCGGTGAATTCACCTTTGACCTCA	2938
Db	112705	TATACCGCTCGACAGATCTTGAATACTGGGAATTCCTCCGGTGAATTCACCTTTGACCTCA	112646
Qy	2939	GTGATGACCAACCTGGTTCTGGCTCTGTGATCTGTTCCCGATGGCTACATGTTGGGAATGCG	2998
Db	112645	GTGATGACCAACCTGGTTCTGGCTCTGTGATCTGTTCCCGATGGCTACATGTTGGGAATGCG	112586
Qy	2999	CCAACTTTTTCACGCTTCGCGATGAAGAACTGGGAAGATCTGACGCTGCTGATTTTCT	3058
Db	112585	CCAACTTTTTCACGCTTCGCGATGAAGAACTGGGAAGATCTGACGCTGCTGATTTTCT	112526
Qy	3059	GTTCACAGGATTTGACCCGAATCCACAGATGATTAATCACTACCAAGCTCTGACCAAT	3118
Db	112535	GTTCACAGGATTTGACCCGAATCCACAGATGATTAATCACTACCAAGCTCTGACCAAT	112466
Qy	3119	GCGGATATGTCGTCCGACCAAGCTTTGAAGAACACACTTCCGCTCTTTGCGAGATTCAGCG	3178
Db	112465	GCGGATATGTCGTCCGCAACACTTGAAGAACACACTTCCGCTCTTTGCGAGATTCAGCG	112406
Qy	3179	AGCTGGAATTTTCGGCCATGATATTTCTACGACCCGAGTTGCACTAAACGGTTCTGATGCTT	3238
Db	112405	AGCTGGAATTTTCGGCCATGATATTTCTACGACCCGAGTTGCACTAAACGGTTCTGATGCTT	112346
Qy	3239	GCGCTGTCGGGCTGGAATGGGGGCTGCCCGCGCAGATGATCAACCAAGTTTGACAGGAAG	3298
Db	112345	GCGCTGTCGGGCTGGAATGGGGGCTGCCCGCGCAGATGATCAACCAAGTTTGACAGGAAG	112286
Qy	3299	GATGGGTGACCTGCTGACTGTGCCCCCGCAGCTTCAATTTGGCAACACACGCGATTAC	3358
Db	112285	GATGGGTGACCTGCTGACTGTGCCCCCGCAGCTTCAATTTGGCAACACACGCGATTAC	112226
Qy	3359	AAGAGCTCCTTCTCCCAAGGGGGAGTGGGGGGTATCAAGTCTGATTAAGGTTCTGAAC	3418
Db	112225	AAGAGCTTCTTCTCCCAAGGGGGAGTCAAGGGTATCAAGTCTGATTAAGGTTCTGAAC	112166
Qy	3419	CTGTCCGAGTAGACATCCGAGGCCAATATTTCCCTCGAGTGGAGTGTGTCCTGTTGTCG	3478
Db	112165	CTGTCCGAGTAGACATCCGAGGCCAATATTTCCCTCGAGTGGAGTGTGTCCTGTTGTCG	112106
Qy	3479	TGGATTCGTGATGCGTATCGTGGCGATGCTGAGGTAAACCTGGCGAATAGTATGGCGG	3538
Db	112105	TGGATTCGTGATGCGTATCGTGGCGATGCTGAGGTAAACCTGGCGAATAGTATGGCGG	112046
Qy	3539	ACGATTAATCAACCCATTTGAGTAACTGACAGTGTATGACAGGTTTCAATTCGCTTTTCGG	3598
Db	112045	ACGATTAATCAACCCATTTGAGTAACTGACAGTGTATGACAGGTTTCAATTCGCTTTCCG	1119878
Qy	3599	GCGTTCAAAAGTGACCTATTGAGAGATTAATTAAGGATTTTGGTGGCGAATTT	3658
Db	111986	GCGTTCAAAAGTGACCTATTGAGAGATTAATTAAGGATTTTGGTGGCGAATTT	1119278
Qy	3659	GTACAAATCTTCGCAAAATCCCTGATGAGACAAATTAACAGGTTTAATTAATTTGTTA	3718
Db	111926	GTACAAATCTTCGCAAAATCCCTGATGAGACAAATTAACAGGTTTAATTAATTTGTTA	1118678
Qy	3719	GCTTTTGAACAAACATTCATGTTCTGAATATTTTGTCTTCCCGTTAAGGAAATTC	3778
Db	111866	GCTTTTGAACAAACATTCATGTTCTGAATATTTTGTCTTCCCGTTAAGGAAATTC	1118078
Qy	3779	ATGAGACATTAAGACCTCGCGCAACGCAATCCGCGGACATTTGGCGGCAAGCAACATT	3838
Db	111806	ATGAGACATTAAGACCTCGCGCAACGCAATCCGCGGACATTTGGCGGCAAGCAACATT	1117478

QY 3839 GTTCGCCGCCGACACTGTGCAAGCGTTTACGCTGTGTCTCAAGACACCAAGATGTC 3898
DB 111746 GTTCGCCGCCGACACTGTGCAAGCGTTTACGCTGTGTCTCAAGACACCAAGATGTC 111687
QY 3899 GATTCGCAAGCTGTGATGATGATCCAGATCTGAAGGACACCTTTGAACCTGCGCGCATG 3958
DB 111686 GATTCGCAAGCTGTGATGATGATCCAGATCTGAAGGACACCTTTGAACCGGATGATG 111627
QY 3959 TTTCAGATCATGTCGCGGCGCAGGCGATGTGATCATGTTTCAAGAACTCATGACGA 4018
DB 111626 TTTCAGATCATGTCGCGGCGCAGGCGATGTGATCATGTTTCAAGAACTCATGACGA 111567
QY 4019 ACCTCCAAAGACATCGCTGTGTCTCAAGACAGCTCAAGATGTTGTGCTAAACGCGC 4078
DB 111566 ACCTCCAAAGACATCGCTGTGTCTCAAGACAGCTCAAGATGTTGTGCTAAACGCGC 111507
QY 4079 AACTGGTTACGCGGTGTGAAAGGATTTGGCGGACATTTGCTCCGCTGATCCATTC 4138
DB 111506 AACTGGTTACGCGGTGTGAAAGGATTTGGCGGACATTTGCTCCGCTGATCCATTC 111447
QY 4139 TTGGTTGTGCGGCTGTGCTCAATGCTATCAACATGTGTTGTCGAGATCTGTTTC 4198
DB 111446 TTGGTTGTGCGGCTGTGCTCAATGCTATCAACATGTGTTGTCGAGATCTGTTTC 111387
QY 4199 GGTTCGCAATCACTGTGAGATGTTCCCTCAGATCAGCGGTGTTCTGATGATCAAC 4258
DB 111386 GGTTCGCAATCACTGTGAGATGTTCCCTCAGATCAGCGGTGTTCTGATGATCAAC 111327
QY 4259 CTGATGCGCATTCGCGCGGTTGCGTCTTGCAGGTGTTGTTGTTTCCGCAACG 4318
DB 111326 CTGATGCGCATTCGCGCGGTTGCGTCTTGCAGGTGTTGTTGTTTCCGCAACG 111267
QY 4319 CGTTTCGATGCGCAATGATGTTCTGCGGCGCGCATTTGATGCGATGTTTCCCAAC 4378
DB 111266 CGTTTCGATGCGCAATGATGTTCTGCGGCGCGCATTTGATGCGATGTTTCCCAAC 111207
QY 4379 CTGTTTAACGCGCTACGATGCGCGCGCACCATGACCGCGCGGCAATGTCATGTGTC 4438
DB 111206 CTGTTTAACGCGCTACGATGCGCGCGCACCATGACCGCGCGGCAATGTCATGTGTC 111147
QY 4439 CTGTTTGGTTGGATGTTGCTCAAGCTGTTACCAAGGACCGTCTTCTGTGCTGTC 4498
DB 111146 CTGTTTGGTTGGATGTTGCTCAAGCTGTTACCAAGGACCGTCTTCTGTGCTGTC 111087
QY 4499 GTCTCTGATTTCTGCGCAACGATGCAAGATTTCTGCAAGGACACTGAGCACTGCA 4558
DB 111086 GTCTCTGATTTCTGCGCAACGATGCAAGATTTCTGCAAGGACACTGAGCACTGCA 111027
QY 4559 GACTTCTGATCAACCCAGTGTGACTGTGCTGCTCAACCGGCTTCTTACGTTTATGCT 4618
DB 111026 GACTTCTGATCAACCCAGTGTGACTGTGCTGCTCAACCGGCTTCTTACGTTTATGCT 110967
QY 4619 ATTGTGCGCAATGCGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4678
DB 110966 ATTGTGCGCAATGCGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 110907
QY 4679 GATTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4738
DB 110906 GATTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 110847
QY 4739 ACTGCTGCGCAACGCTTCTCCGCAATTTAGCTGAGCTGTTCAACCAAGGCTGATCC 4798
DB 110846 ACTGCTGCGCAACGCTTCTCCGCAATTTAGCTGAGCTGTTCAACCAAGGCTGATCC 110787
QY 4799 TTTCATTTGCGCAACCGCATTCATGCGCAATTCGCGCAAGGCTGCGCATGTTGCGAGTG 4858
DB 110786 TTTCATTTGCGCAACCGCATTCATGCGCAATTCGCGCAAGGCTGCGCATGTTGCGAGTG 110727
QY 4859 TTCTTCTAGCGAAGATGAAAAAGCTCAAGGCGCTTGCAAGTGTCTTCCGCT 4918
DB 110726 TTCTTCTAGCGAAGATGAAAAAGCTCAAGGCGCTTGCAAGTGTCTTCCGCT 110667

QY 4919 GTTCTTGGATTTACAGAGCTGCGATCTTCCGTTGGAACCTTGGCTGCGGCTGTC 4978
DB 110666 GTTCTTGGATTTACAGAGCTGCGATCTTCCGTTGGAACCTTGGCTGCGGCTGTC 110607
QY 4979 TACATTTGATTCGATTCGCGAGCTATCGGTGCGCTTGTGATTCATCTTGTATCAAG 5038
DB 110606 TACATTTGATTCGATTCGCGAGCTATCGGTGCGCTTGTGATTCATCTTGTATCAAG 110547
QY 5039 GCAATTTGCGGTGCGGCTGCGAGTTCCTTGGGTGTTGTTTATGATGCTTCAGATATG 5098
DB 110546 GCAATTTGCGGTGCGGCTGCGAGTTCCTTGGGTGTTGTTTATGATGCTTCAGATATG 110487
QY 5099 GTTCATGTTCTGTTGCGGCTGCTTGTGCTTGTGATTCGATTCGATTCGATTCGATTC 5158
DB 110486 GTTCATGTTCTGTTGCGGCTGCTTGTGCTTGTGATTCGATTCGATTCGATTCGATTC 110427
QY 5159 TATGCGCTTATCTGTTGCTGCGCAACGCGAGCATTTGATTCAGATGCAACGCTCTCA 5218
DB 110426 TATGCGCTTATCTGTTGCTGCGCAACGCGAGCATTTGATTCAGATGCAACGCTCTCA 110367
QY 5219 GTGCTTGCAGAGCAACCAAGCCGAGAGACCCGAGATTTTCAACGATTC 5278
DB 110366 GTGCTTGCAGAGCAACCAAGCCGAGAGACCCGAGATTTTCAACGATTC 110307
QY 5279 ACCATTCATTCAGGACCTTTGATCCGCTGATGCTATGCACTGAGAGCGTCCGATGTC 5338
DB 110306 ACCATTCATTCAGGACCTTTGATCCGCTGATGCTATGCACTGAGAGCGTCCGATGTC 110247
QY 5339 ATGTTTGCAGGAGAACTTGTGCTCAGGTGTTGCGATGCTGCTCCCAACGAGGAGCTG 5398
DB 110246 ATGTTTGCAGGAGAACTTGTGCTCAGGTGTTGCGATGCTGCTCCCAACGAGGAGCTG 110187
QY 5399 GTTTCACAGTAGAGCGGAAAGATGCTGCTGCTGCTTCCATCTGTCAGCTTTCGATGTC 5458
DB 110186 GTTTCACAGTAGAGCGGAAAGATGCTGCTGCTGCTTCCATCTGTCAGCTTTCGATGTC 110127
QY 5459 CGCATTAAGCTGAGAGATGCTTCCATGATGATCTTGAATGCTGATGCTTTCGACACC 5518
DB 110126 CGCATTAAGCTGAGAGATGCTTCCATGATGATCTTGAATGCTGATGCTTTCGACACC 110067
QY 5519 GTTAACTCAACGCGAGCACTTTAACCGCTGAGAGAGAGCGAGTGAAGTCAAGCA 5578
DB 110066 GTTAACTCAACGCGAGCACTTTAACCGCTGAGAGAGAGCGAGTGAAGTCAAGCA 110007
QY 5579 GGGAGCTGCTGTGTAATTCGATATGATGATGATGATGATGATGATGATGATGATGAT 5638
DB 110006 GGGAGCTGCTGTGTAATTCGATATGATGATGATGATGATGATGATGATGATGATGAT 109947
QY 5639 AGCGGATTTGTTTTCGATTAAGAAACCGGACCTGTAACACTTACGCTTGTGCGC 5698
DB 109946 AGCGGATTTGTTTTCGATTAAGAAACCGGACCTGTAACACTTACGCTTGTGCGC 109887
QY 5699 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCGACAACA 5758
DB 109886 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCGACAACA 109827
QY 5759 CCATAAGTTGAACCTTGAATGTTGCAACAGGTTTGAACCTTGAAGGAGCTGACCTTACGC 5818
DB 109826 CCATAAGTTGAACCTTGAATGTTGCAACAGGTTTGAACCTTGAAGGAGCTGACCTTACGC 109767
QY 5819 ATCTTGAACACCGGATCCGCTGACCTTGAATTTTAAACCTGTTCAACAGGATGATGCG 5878
DB 109766 ATCTTGAACACCGGATCCGCTGACCTTGAATTTTAAACCTGTTCAACAGGATGATGCG 109707
QY 5879 TCGGTGTCCTGTTGTTGTCGCAACCCGCAATCTTCAACCCCACTTGAAGATGTTGTTCA 5938
DB 109706 TCGGTGTCCTGTTGTTGTCGCAACCCGCAATCTTCAACCCCACTTGAAGATGTTGTTCA 109647
QY 5939 GCAATGAGCTTATTTTGTGCGGCTG 5967
DB 109646 GCAATGAGCTTATTTTGTGCGGCTG 109618

```

RESULT 5
E11760
LOCUS      E11760      6911 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION Base sequence of sucrase gene.
ACCESSION  E11760
VERSION    E11760.1 GI:22025396
KEYWORDS   JP 1996196280-A/1.
SOURCE     Corynebacterium glutamicum
            Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE  1 [bases 1 to 6911]
            Sugimoto, M., Otona, K., Nagase, K., Tsuchiya, M., Matsui, Y.,
            Yoshihara, Y. and Nakamatsu, M.
            SUCROSE GENE ORIGINATED FROM CORYNEFORM BACTERIUM AND ITS
            PATENT: JP 1996196280-A 1 06-AUG-1996;
            AJINOMOTO CO INC
COMMENT    OS Brevibacterium lactofermentum
            PN JP 1996196280-A/1
            PD 06-AUG-1996
            PE 30-JAN-1995 JP 1995012361
            PI SUGIMOTO MASAKAZU, OTONA KIYOKO, NAGASE KAZUO, TSUCHIYA
            MAKOTO, PI MATSUI YUTAKA, YOSHIHARA YASUHIKO, NAKAMATSU WATARU PC
            C12N15/09, C12N1/21, C12P13/06//C12N9/24, (C12N1/21, C12R1:15); CC
            strandedness: Double;
            CC topology: Linear;
            FH Location/Qualifiers
            FT source
            FT 1..6911
            FT /organism='Brevibacterium lactofermentum' FT
            FT /strain='ATCC13869'
            FT /plasmid='pSSM30BS'
            FT 342..1505
            FT /product='A protein having high homology to FT
            CDS /product='N-acetylglucosamine-6-phosphate deacetylase' FT
            FT 2338..3609
            FT /product='Sucrase'
            FT 4438..5358
            FT /product='A protein having high homology to FT
            FT CDS /product='A protein having high homology to FT
            FT 5570..6577
            FT /product='A protein having homology to FT
            FT CDS /product='phospho-N-acetylmuramoyl pentapeptide FT
            FT 1..6911
            FT /organism='Corynebacterium glutamicum'
            FT /mol_type='genomic DNA'
            FT /db_xref='taxon:1718'
FEATURES
     source
     location/Qualifiers
     1..6911
     /organism="Corynebacterium glutamicum"
     /mol_type="genomic DNA"
     /db_xref="taxon:1718"
ORIGIN
Query Match      61.8%; Score 3687; DB 6; Length 6911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 241 CACACGAGCGCTCAAGTCCGAGATTAATTAACCTAGATCCGTAGACATTAAGATCATTA 300
DB 241 CAAACGAGCGGCTCAAGTCCGAGATTAATTAACCTAGATCCGTAGACATTAAGATCATTA 300
QY 301 CGTCTATGCTTGTGGAAGAACCAATACTCAGAAAAGATGGCAGAAATGGTGATC 360
DB 301 GGTCTATGCTTGTGGAAGAACCAATACTCAGAAAAGATGGCAGAAATGGTGATC 360
QY 361 ATCAAGAAATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
DB 361 ATCAAGAAATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
QY 421 GGGATATGATGCTTTTCCATCTCCGAAACCGCATATATCAAGGAATCTCTGAGAAC 480
DB 421 GGGATATGATGCTTTTCCATCTCCGAAACCGCATATATCAAGGAATCTCTGAGAAC 480
QY 481 CAGACCTAAATAACGAGAGATTCACACCCGAACTCCACAGATTTGTTCCGGTTTATG 540
DB 481 CAGACCTAAATAACGAGAGATTCACACCCGAACTCCACAGATTTGTTCCGGTTTATG 540
QY 541 ATCTTCATATACAGGTGGAACCGGTGCGCTTTCTTACCGGAAACGAGACCAAGCGCA 600
DB 541 ATCTTCATATACAGGTGGAACCGGTGCGCTTTCTTACCGGAAACGAGACCAAGCGCA 600
QY 601 GGAACACCGCGCAGATACCGCGAACAATGSCAGACCGCATGTTGCCAAGCATGGTTT 660
DB 601 GGAACACCGCGCAGATACCGCGAACAATGSCAGACCGCATGTTGCCAAGCATGGTTT 660
QY 661 CGGCGCGCGCTGAGCAGCTGGCAGCGCAGGTGGAACCTTTATCCCTTTGTGTAAGAG 720
DB 661 CGGCGCGCGCTGAGCAGCTGGCAGCGCAGGTGGAACCTTTATCCCTTTGTGTAAGAG 720
QY 721 TCTGCTGTGCGCATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 780
DB 721 TCTGCTGTGCGCATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 780
QY 781 AAAACCCGAGTTTCAATTTTCCGCGCAACCCAGAGATCTTCCCGGGTGAATCCATGCGG 840
DB 781 AAAACCCGAGTTTCAATTTTCCGCGCAACCCAGAGATCTTCCCGGGTGAATCCATGCGG 840
QY 841 GAAAGGTTGATCAATATGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 900
DB 841 GAAAGGTTGATCAATATGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 900
QY 901 TCGATCTTGTGCGCAGCGCAGCATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
DB 901 TCGATCTTGTGCGCAGCGCAGCATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
QY 961 ATACCACTACAGAGCGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1020
DB 961 ATACCACTACAGAGCGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1020
QY 1021 ATTGTTCAATGCGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1080
DB 1021 ATTGTTCAATGCGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1080
QY 1081 TTGCTGCGCGCAGCGCGCGGAGCGCATATGTTGATGATGATGCGCGAGCGCGCTGATTTGG 1140
DB 1081 TTGCTGCGCGCAGCGCGCGGAGCGCATATGTTGATGATGATGCGCGAGCGCGCTGATTTGG 1140
QY 1141 CCGATGGAACGGTGAATGATCTGTTCCAAACAGCGCTTTTATCATCAGGACCGCATGG 1200
DB 1141 CCGATGGAACGGTGAATGATCTGTTCCAAACAGCGCTTTTATCATCAGGACCGCATGG 1200
QY 1201 AAGCGCGGAGATGCGAGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB 1201 AAGCGCGGAGATGCGAGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 ATGAGATGCGCGCTGCGCATGCGCGCGCGCATATGCGCGCGCGCGCATATGCGCGCGCGCG 1320
DB 1261 ATGAGATGCGCGCTGCGCATGCGCGCGCGCATATGCGCGCGCGCGCATATGCGCGCGCGCG 1320
QY 1321 GTCAAGTTGTGCAACAGCGTGGCAGGGGTATGACGTTATGACGCGACCTTCACACT 1380

```



```
Db 1321 GTGAGTTCGTGCAACGTCGCGAGGGGRTGAGCGTTATCGACGGGACCCCTTCAACCT 1380
Qy 1381 CAACGTCGCGCTAAATTTCTCGGTCCTGGCGATCAAGAAATCGTAAATCCAACTCG 1440
Db 1381 CAACGTCGCGCTAAATTTCTCGGTCCTGGCGATCAAGAAATCGTAAATCCAACTCG 1440
Qy 1441 CAAATTTTGTGCTTTGACTGCAAAAGGCGAGTGCAGAAAGGTCCATTTAGTCAACG 1500
Db 1441 CAAATTTTGTGCTTTGACTGCAAAAGGCGAGTGCAGAAAGGTCCATTTAGTCAACG 1500
Qy 1501 TACTTTAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
Db 1501 TACTTTAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
Qy 1561 TCTGCAAGACGAGCAAGAGTGGCAAGCAAGTTCAGTCTTAATCGCACTCTTGGCA 1620
Db 1561 TCTGCAAGACGAGCAAGAGTGGCAAGCAAGTTCAGTCTTAATCGCACTCTTGGCA 1620
Qy 1621 ACAAGGTTGGAACCTTGGGGCTTGCAACAGATCTTCACTGAGTACTTACCAAGAGC 1680
Db 1621 ACAAGGTTGGAACCTTGGGGCTTGCAACAGATCTTCACTGAGTACTTACCAAGAGC 1680
Qy 1681 TCATTCGCAATGATGAGAGTGGGGAAGTGTCACTTCAAGAACTGCAAGGACTTCTTGG 1740
Db 1681 TCATTCGCAATGATGAGAGTGGGGAAGTGTCACTTCAAGAACTGCAAGGACTTCTTGG 1740
Qy 1741 ATGAAATAGCTGGAGCTTAACCCGTGACGATGAAACAGCTATTTAAACCATTCGCAAG 1800
Db 1741 ATGAAATAGCTGGAGCTTAACCCGTGACGATGAAACAGCTATTTAAACCATTCGCAAG 1800
Qy 1801 AGTTCACTGACCACTGCACTGTGTGATGAAAGGTCTTACAGCCAGATGTGTGCAAC 1860
Db 1801 AGTTCACTGACCACTGCACTGTGTGATGAAAGGTCTTACAGCCAGATGTGTGCAAC 1860
Qy 1801 AGTTCACTGACCACTGCACTGTGTGATGAAAGGTCTTACAGCCAGATGTGTGCAAC 1860
Db 1801 AGTTCACTGACCACTGCACTGTGTGATGAAAGGTCTTACAGCCAGATGTGTGCAAC 1860
Qy 1861 CTGATCCATACGAAAGAGCTGACAGATGATGAGCAAGATCGCTGAGATCCGTTGAAG 1920
Db 1861 CTGATCCATACGAAAGAGCTGACAGATGATGAGCAAGATCGCTGAGATCCGTTGAAG 1920
Qy 1921 TTCAAAATCCTTGGCATCGGCGGAAACGGCACTCGTTTCACTTGAACATCATCTTCTGT 1980
Db 1921 TTCAAAATCCTTGGCATCGGCGGAAACGGCACTCGTTTCACTTGAACATCATCTTCTGT 1980
Qy 1981 GTCAAGACTGACAAAGGTCCAGGCGCTGCACTTAAACCTGTGAGAGCAACCGCTGAT 2040
Db 1981 GTCAAGACTGACAAAGGTCCAGGCGCTGCACTTAAACCTGTGAGAGCAACCGCTGAT 2040
Qy 2041 CTTCAACACCATGAGAGGTCCCAACCAACGCGCTCACCCAGGGTTGGGCACTTGTCT 2100
Db 2041 CTTCAACACCATGAGAGGTCCCAACCAACGCGCTCACCCAGGGTTGGGCACTTGTCT 2100
Qy 2101 CCGCGGCAAAACATCGTGTGTGGCAACTGTGTGAAGAAAGCCGACGCCATTCGCGG 2160
Db 2101 CCGCGGCAAAACATCGTGTGTGGCAACTGTGTGAAGAAAGCCGACGCCATTCGCGG 2160
Qy 2161 AACTGTGGAAGGCGCCAGTACTGTCTTGTGCCAGGTTCATCTCTGTAGTGAACAAT 2220
Db 2161 AACTGTGGAAGGCGCCAGTACTGTCTTGTGCCAGGTTCATCTCTGTAGTGAACAAT 2220
Qy 2161 AACTGTGGAAGGCGCCAGTACTGTCTTGTGCCAGGTTCATCTCTGTAGTGAACAAT 2220
Db 2161 AACTGTGGAAGGCGCCAGTACTGTCTTGTGCCAGGTTCATCTCTGTAGTGAACAAT 2220
Qy 2221 GGCACATCATGTTGGATGAGAGCAGAGTATCAAGCTGGAAGAAAGCTGATCACTAAC 2280
Db 2221 GGCACATCATGTTGGATGAGAGCAGAGTATCAAGCTGGAAGAAAGCTGATCACTAAC 2280
Qy 2281 TCTCATGAGCAATTAAGCTGCGTGAAGAAACAAAGAAAGTACTGTGTGGGCTATG 2340
Db 2281 TCTCATGAGCAATTAAGCTGCGTGAAGAAACAAAGAAAGTACTGTGTGGGCTATG 2340
Qy 2341 CACACGAACCTTCCAGTTTGGCGCTGGGTACATGTGATCTCTCCGAGGGCAGGCTC 2400
Db 2341 CACACGAACCTTCCAGTTTGGCGCTGGGTACATGTGATCTCTCCGAGGGCAGGCTC 2400
Qy 2401 AATGATCCCAAGCAATGATACGTGATGAGATACCTCCAGTCACTACTACAGACGAT 2460
Db 2401 AATGATCCCAAGCAATGATACGTGATGAGATACCTCCAGTCACTACTACTACAGACGAT 2460
Db 2401 AATGATCCCAAGCAATGATACGTGATGAGATACCTCCAGTCACTACTACTACTACAGACGAT 2460
Qy 2461 CCAAGTTTCCCTTGGCACCAAGGCGACCGGCTGGGCTCACACCAACGCGCTTGACC 2520
Db 2461 CCAAGTTTCCCTTGGCACCAAGGCGACCGGCTGGGCTCACACCAACGCGCTTGACC 2520
Qy 2521 GGAAGCGAGGATGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Db 2521 GGAAGCGAGGATGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Qy 2581 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
Db 2581 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
Qy 2641 TACACCGGCACTTAAATTTGACGAAAGGCGCGGCAACCAAACTTGTGCAAGTC 2700
Db 2641 TACACCGGCACTTAAATTTGACGAAAGGCGCGGCAACCAAACTTGTGCAAGTC 2700
Qy 2701 GAGGACCCCACTGGGCTGATGGGCGGCACTTATCGCGCTTGAATGCGCTTATC 2760
Db 2701 GAGGACCCCACTGGGCTGATGGGCGGCACTTATCGCGCTTGAATGCGCTTATC 2760
Qy 2761 GACGAGCCGCGAGGCTTTCACACCCATTAACCGGATCCGATGATGATGATGATGATGAT 2820
Db 2761 GACGAGCCGCGAGGCTTTCACACCCATTAACCGGATCCGATGATGATGATGATGATGAT 2820
Qy 2821 GATGTTGGAACATGTTTGGGCGCAACGCGAAACCTGACCGGTGACGCTTCTA 2880
Db 2821 GATGTTGGAACATGTTTGGGCGCAACGCGAAACCTGACCGGTGACGCTTCTA 2880
Qy 2881 TACCGCTGACGATCTTGAAGAACTGGGAAATCTCGGGAAATGACCTTTGACCTGCT 2940
Db 2881 TACCGCTGACGATCTTGAAGAACTGGGAAATCTCGGGAAATGACCTTTGACCTGCT 2940
Qy 2941 GATGCAACACTGTTCTGCTCTGATCTGTTCCGATGATGATGATGATGATGATGATGAT 3000
Db 2941 GATGCAACACTGTTCTGCTCTGATCTGTTCCGATGATGATGATGATGATGATGATGAT 3000
Qy 3001 AACCTTTTTCGCTTCCGATGAGAAACCTGGCGAAGATCTGACGCTGATTTTCTGT 3060
Db 3001 AACCTTTTTCGCTTCCGATGAGAAACCTGGCGAAGATCTGACGCTGATTTTCTGT 3060
Qy 3061 CCAAGAGATGAGACCGAATCCAGATGAGATGATGATGATGATGATGATGATGATGAT 3120
Db 3061 CCAAGAGATGAGACCGAATCCAGATGAGATGATGATGATGATGATGATGATGATGAT 3120
Qy 3121 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
Db 3121 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
Qy 3181 TCGATTTTGGGCAATGATTTTACGACCGGAGTTGAGTAAACGTTTCTGATGCTGG 3240
Db 3181 TCGATTTTGGGCAATGATTTTACGACCGGAGTTGAGTAAACGTTTCTGATGCTGG 3240
Qy 3241 CTGCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Db 3241 CTGCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Qy 3301 TGGGTGCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Db 3301 TGGGTGCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Qy 3361 GAGCTCTTCTTCCAGAGGAGGATGAGGAGGATGATGATGATGATGATGATGATGATGAT 3420
Db 3361 GAGCTCTTCTTCCAGAGGAGGATGAGGAGGATGATGATGATGATGATGATGATGATGAT 3420
Qy 3421 GTCCAGTGAAGATCCGAGGCAATTTTCCCTGAGTGGAGAGTGTCCGTTTGTCTGG 3480
Db 3421 GTCCAGTGAAGATCCGAGGCAATTTTCCCTGAGTGGAGAGTGTCCGTTTGTCTGG 3480
Qy 3481 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
Db 3481 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
```

QY 3541 GATATATACGCCATTGAGATATCTGACAGGTGATGACAGCTTTCTATGCTTTTCCGGGC 3600
|
|
|
Db 3541 GATATATACGCCATTGAGATATCTGACAGGTGATGACAGCTTTCTATGCTTTTCCGGGC 3600
|
|
|
QY 3601 CTTCAGAGGTGACATATTTAGAGATTAAGTCATATAAAGGCTCTTTTGGCGAATTGT 3660
|
|
|
Db 3601 CTTCAGAGGTGACATATTTAGAGATTAAGTCATATAAAGGCTCTTTTGGCGAATTGT 3660
|
|
|
QY 3661 ACAAACTCTGCGCAAAATCCCTTGATC 3687
|
|
|
Db 3661 ACAAACTCTGCGCAAAATCCCTTGATC 3687
|
|
|
RESULT 6
126124 6911 bp DNA linear PAT 07-OCT-1996
126124
LOCUS Sequence 4 from patent US 556776.
DEFINITION
ACCESSION 126124
VERSION 126124.1 GI:1605994
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6911)
AUTHORS Teuchiya,M. and Miwa,K.
TITLE Sucrase gene derived from *coxyneiform* bacteria
JOURNAL Patent: US 556776-A 4 17-SEP-1996;
FEATURES
source Location/Qualifiers
1..6911
/organism="unknown"
/mol_type="unasigned DNA"

Query Match 61.8%; Score 3687; DB 6; Length 6911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCCGTGACGCCACCATTTGATGTCGTGTCACGAGCTTCCGAGCTTCTACATCT 60
|
|
|
Db 1 AGTCCGTGACGCCACCATTTGATGTCGTGTCACGAGCTTCCGAGCTTCTACATCT 60
|
|
|
QY 61 AGCTCCCGTCGCGGTGAGTGGGGTCATTACGGGTGGATCACCGCGTGAAGTTGGC 120
|
|
|
Db 61 AGCTCCCGTCGCGGTGAGTGGGGTCATTACGGGTGGATCACCGCGTGAAGTTGGC 120
|
|
|
QY 121 GAACCCATGCTCTCTGTTGGGTTGAGGGAACGAGTGGCGGTGAGAACTTTTCAAGTG 180
|
|
|
Db 121 GAACCCATGCTCTCTGTTGGGTTGAGGGAACGAGTGGCGGTGAGAACTTTTCAAGTG 180
|
|
|
QY 181 TCTGCAAGTTTAAAGTATGATCATCATCTTGAAGGCTGAGATTAATCAATAGCCTG 240
|
|
|
Db 181 TCTGCAAGTTTAAAGTATGATCATCATCTTGAAGGCTGAGATTAATCAATAGCCTG 240
|
|
|
QY 241 CAACAGCAGGCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAGCATATAGACATCA 300
|
|
|
Db 241 CAACAGCAGGCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAGCATATAGACATCA 300
|
|
|
QY 301 CGTCTATGCTTGTGAGAGAAACCAATTAACCTGAGAAAGATGGCAGAAAGTGGCAT 360
|
|
|
Db 301 CGTCTATGCTTGTGAGAGAAACCAATTAACCTGAGAAAGATGGCAGAAAGTGGCAT 360
|
|
|
QY 361 ATCAAGAAATGCAAGTCAAGCATTAATAAATTAAGGAGAAATGTTCCCGCCCTG 420
|
|
|
Db 361 ATCAAGAAATGCAAGTCAAGCATTAATAAATTAAGGAGAAATGTTCCCGCCCTG 420
|
|
|
QY 421 GGGTGTGATGAGCTTTCTCAACTGAAAACGCGCATCATCGGAACCTCTGAGAAC 480
|
|
|
Db 421 GGGTGTGATGAGCTTTCTCAACTGAAAACGCGCATCATCGGAACCTCTGAGAAC 480
|
|
|
QY 481 CAGCACCTTAAACGAGATTCACCCGAACTCCCAAGTGTTCCTGGGTTTATG 540
|
|
|
Db 481 CAGCACCTTAAACGAGATTCACCCGAACTCCCAAGTGTTCCTGGGTTTATG 540
|
|
|

QY 541 ATCTTCAATATCAGGTTGGAACGAGTGGCGGCTTCTACAGGGAACGAGGACGAGCGA 600
|
|
|
Db 541 ATCTTCAATATCAGGTTGGAACGAGTGGCGGCTTCTACAGGGAACGAGGACGAGCGA 600
|
|
|
QY 601 GGAACACCGCGCAGTATCACCGCAACATGGAACGACCGTATGTTGCCAAGATGTTT 660
|
|
|
Db 601 GGAACACCGCGCAGTATCACCGCAACATGGAACGACCGTATGTTGCCAAGATGTTT 660
|
|
|
QY 661 CGGCGCGCGCTGACGACATGCGACGCGAGGTGAAACCTTATTCCTTGTGTGAAGG 720
|
|
|
Db 661 CGGCGCGCGCTGACGACATGCGACGCGAGGTGAAACCTTATTCCTTGTGTGAAGG 720
|
|
|
QY 721 TCTGCTGTGCGGATTCACCTGAGGCGCTTTCATCAACGACGCGGTGTGCTGCTC 780
|
|
|
Db 721 TCTGCTGTGCGGATTCACCTGAGGCGCTTTCATCAACGACGCGGTGTGCTGCTC 780
|
|
|
QY 781 AAAACCGGATTTTATTTTCCCGCAACCCCAAGATCTTCCCGGTGATCCATGCGG 840
|
|
|
Db 781 AAAACCGGATTTTATTTTCCCGCAACCCCAAGATCTTCCCGGTGATCCATGCGG 840
|
|
|
QY 841 GAAAAGTTGGATCAATTCATCAAGTACAGTACGCGCGGAACTGACATCTTTTGAAGTTC 900
|
|
|
Db 841 GAAAAGTTGGATCAATTCATCAAGTACAGTACGCGCGGAACTGACATCTTTTGAAGTTC 900
|
|
|
QY 901 TCGATCTGCGGACGCGCACATCATATGCTTCCCTGCGGACACATGATGCAATTTTG 960
|
|
|
Db 901 TCGATCTGCGGACGCGCACATCATATGCTTCCCTGCGGACACATGATGCAATTTTG 960
|
|
|
QY 961 ATACCACTACGAGCGCAATTTCTTGGTAAAGAAAATGTGACGCTACGCGCTACGC 1020
|
|
|
Db 961 ATACCACTACGAGCGCAATTTCTTGGTAAAGAAAATGTGACGCTACGCGCTACGC 1020
|
|
|
QY 1021 ATTTGTTAATGCGATGCTCCGCTGATCATATAGGCTCCCGGAGAGTGGCGCTTGC 1080
|
|
|
Db 1021 ATTTGTTAATGCGATGCTCCGCTGATCATATAGGCTCCCGGAGAGTGGCGCTTGC 1080
|
|
|
QY 1081 TTGCTGCGGACGTCGCGGAGCGCATATGTTGATGATGCGGACGCGGTGATTTGG 1140
|
|
|
Db 1081 TTGCTGCGGACGTCGCGGAGCGCATATGTTGATGATGCGGACGCGGTGATTTGG 1140
|
|
|
QY 1141 CCGATGAACGCTGATCTGCTGTCACAAACGCTTTTTCATCAACGACCGCATG 1200
|
|
|
Db 1141 CCGATGAACGCTGATCTGCTGTCACAAACGCTTTTTCATCAACGACCGCATG 1200
|
|
|
QY 1201 AAGCGCGGAATGCGAAGGTGATCATTTTGGGCTTGAAGTCAACCGTACCG 1260
|
|
|
Db 1201 AAGCGCGGAATGCGAAGGTGATCATTTTGGGCTTGAAGTCAACCGTACCG 1260
|
|
|
QY 1261 ATGAGTGGCCGCTGTCGCGATGGCGCGCATGCGCGGGGACACGACATAGCGA 1320
|
|
|
Db 1261 ATGAGTGGCCGCTGTCGCGATGGCGCGCATGCGCGGGGACACGACATAGCGA 1320
|
|
|
QY 1321 GTCAAGTGTGACCAACGTCGCGAGGGTATGACGCTTATGAGCGGACCTTCCACACT 1380
|
|
|
Db 1321 GTCAAGTGTGACCAACGTCGCGAGGGTATGACGCTTATGAGCGGACCTTCCACACT 1380
|
|
|
QY 1381 CAACCGTGGCGCTAAATTTCTCGGTCTTGGGATCAAGAAATGCTAAATCCAACCTG 1440
|
|
|
Db 1381 CAACCGTGGCGCTAAATTTCTCGGTCTTGGGATCAAGAAATGCTAAATCCAACCTG 1440
|
|
|
QY 1441 CAATTTTGTGCTTTGACTCAACGCGCAGGTGCAAAAGTCCATTTAGATCAAG 1500
|
|
|
Db 1441 CAATTTTGTGCTTTGACTCAACGCGCAGGTGCAAAAGTCCATTTAGATCAAG 1500
|
|
|
QY 1501 TACTTTAAGTACGATTAATCTATCTGATTTTAAAGAGTCCCACTGAAATACATA 1560
|
|
|
Db 1501 TACTTTAAGTACGATTAATCTATCTGATTTTAAAGAGTCCCACTGAAATACATA 1560
|
|
|
QY 1561 TCTGCAAGAGAGAGAAAGTGGCAGAGTTCAGTCTATGACACCTTGGCA 1620
|
|
|
Db 1561 TCTGCAAGAGAGAGAAAGTGGCAGAGTTCAGTCTATGACACCTTGGCA 1620
|
|
|

QY	1621	ACAAAGGTGGAA	CC	TTGGGGCTTGGAA	CAGAA	TCTCA	CACTAGAT	CTTACCA	AGC	1680
Db	1621	ACAAAGGTGAA	CC	TTGGGGCTTGGAA	CAGAA	TCTCA	CACTAGAT	CTTACCA	AGC	1680
QY	1681	TCAATTCGATGATGA	TAAAGCTGGGGAAG	TGTCATTTCA	TAAAGAC	CTGCAAGG	CATTCTGTGG		1740	
Db	1681	TCATTCGATGATGATGA	TAAAGCTGGGGAAG	TGTCATTTCA	TAAAGAC	CTGCAAGG	CATTCTGTGG		1740	
QY	1741	ATGAATACGTGGGA	CTAA	CCCGTGA	CGATGAAA	CAGCTA	CTTTAAAC	CAATTGCAAG	1800	
Db	1741	ATGAATACGTGGGA	CTAA	CCCGTGA	CGATGAAA	CAGCTA	CTTTAAAC	CAATTGCAAG	1800	
QY	1801	AGTTCACTGA	CCACATTCGACAT	CGTTGATGA	GAAGGTCT	CAAGCC	CAGATGTC	AAAC	1860	
Db	1801	AGTTCACTGA	CCACATTCGACAT	CGTTGATGA	GAAGGTCT	CAAGCC	CAGATGTC	AAAC	1860	
QY	1861	CTGATCCATPACGA	ACAGCTGACAGTATG	AGGCGAA	ATAATGCTG	ACAAATCCG	TGAA	1920		
Db	1861	CTGATCCATPACGA	ACAGCTGACAGTATG	AGGCGAA	ATAATGCTG	ACAAATCCG	TGAA	1920		
QY	1921	TTCAAAATCCTTGG	CAATCGCGGAAA	CGGCA	CATCGCTTTCA	TGAAAC	CATCATCTCT	1980		
Db	1921	TTCAAAATCCTTGG	CAATCGCGGAAA	CGGCA	CATCGCTTTCA	TGAAAC	CATCATCTCT	1980		
QY	1981	GTCAAGACTGACAA	AAAGTTC	CAAGCGCTG	CACTTAAAC	CTGGA	AGCAAC	CGTCGAT	2040	
Db	1981	GTCAAGACTGACAA	AAAGTTC	CAAGCGCTG	CACTTAAAC	CTGGA	AGCAAC	CGTCGAT	2040	
QY	2041	CTTCAACACATCG	AAAGAGTCCCA	CCCA	CGCGCTCA	CCAGGTTTGGG	CACTTGT	2100		
Db	2041	CTTCAACACATCG	AAAGAGTCCCA	CCCA	CGCGCTCA	CCAGGTTTGGG	CACTTGT	2100		
QY	2101	CCGCGCGGAAA	CATCGTGTGTTGG	TGGCAACTG	TGTAAGAAAG	CCGAC	GCATCCGCG	2160		
Db	2101	CCGCGCGGAAA	CATCGTGTGTTGG	TGGCAACTG	TGTAAGAAAG	CCGAC	GCATCCGCG	2160		
QY	2161	AACTGTGAAGAG	CCCACTGATCTG	CTTCTTCC	CCAGGTTCCAT	CTGTAGAT	GCACACAT	2220		
Db	2161	AACTGTGAAGAG	CCCACTGATCTG	CTTCTTCC	CCAGGTTCCAT	CTGTAGAT	GCACACAT	2220		
QY	2221	GCCACCATCATCG	TTGGATGAAAG	CGAGTATCCAA	AGCTGAAAAC	CGTATCA	CTACCG	2280		
Db	2221	GCCACCATCATCG	TTGGATGAAAG	CGAGTATCCAA	AGCTGAAAAC	CGTATCA	CTACCG	2280		
QY	2281	TCTCATGAGCA	ATTAAAGCTG	CGCTAGAAA	CAAAAAGTA	ACTGTG	GGGCTATG	2340		
Db	2281	TCTCATGAGCA	ATTAAAGCTG	CGCTAGAAA	CAAAAAGTA	ACTGTG	GGGCTATG	2340		
QY	2341	CACACAGAA	CTTTCAG	TTTTCG	CGCTGATCA	TGTA	CTCTCG	CAAGGCTC	2400	
Db	2341	CACACAGAA	CTTTCAG	TTTTCG	CGCTGATCA	TGTA	CTCTCG	CAAGGCTC	2400	
QY	2401	AATGATCCCA	ACGAAATGATCG	ATGGAATAC	CTTCCAG	GTCTACTAC	ACAGAT	2460		
Db	2401	AATGATCCCA	ACGAAATGATCG	ATGGAATAC	CTTCCAG	GTCTACTAC	ACAGAT	2460		
QY	2461	CCAGGTTTCC	CTTTCGACAA	AGGCGAC	CGGCTG	GCACACA	CAAGCG	CTTATC	2520	
Db	2461	CCAGGTTTCC	CTTTCGACAA	AGGCGAC	CGGCTG	GCACACA	CAAGCG	CTTATC	2520	
QY	2521	GGAACCGAC	GATGACATG	CGACGCA	CTG	CCGACG	CTTTAC	CCGGAATG	CTAT	2580
Db	2521	GGAACCGAC	GATGACATG	CGACGCA	CTG	CCGACG	CTTTAC	CCGGAATG	CTAT	2580
QY	2581	GACCTGATG	ATGATTTCCG	GTGAGAC	CGGTATTTA	TGACG	CAACATTTAA	CTTTTC	2640	
Db	2581	GACCTGATG	ATGATTTCCG	GTGAGAC	CGGTATTTA	TGACG	CAACATTTAA	CTTTTC	2640	
QY	2641	TACACCGGCA	ACTTAAATTTGA	CGGAAAG	CGCGGCA	CCCAAAA	CTTTGTG	AGATC	2700	
Db	2641	TACACCGGCA	ACTTAAATTTGA	CGGAAAG	CGCGGCA	CCCAAAA	CTTTGTG	AGATC	2700	
QY	2701	GAGGACCCAA	CTGGGCTGATG	GGGCGGCA	TTTCATCG	CGCGTGTG	CCCTAA	AAATCCG	CTTATC	2760

Db	2701	GAGGACCCAACTGGGCTGATGGGGGCAATTCATGCGCTTGCTTAATAAAATCCGCTTATC	2760
Qy	2761	GACGGACCCGCGACGGGTTTCAACCCCATTAACCGCATCCCATGATCAGCCGTGATAGT	2820
Db	2761	GACGGACCCGCGACGGGTTTCAACCCCATTAACCGCATCCCATGATCAGCCGTGATAGT	2820
Qy	2821	GATGTTTGAACATGCTTCTGGGGGCCAACGCGAAAACCTCACCGGTGACGGGTTCTA	2880
Db	2821	GATGTTTGAACATGCTTCTGGGGGCCAACGCGAAAACCTCACCGGTGACGGGTTCTA	2880
Qy	2881	TACGCTGACAGATCTTGAAGAACTGGGGAATTCCTCGGTGAATCACCCTTTGACTAGT	2940
Db	2881	TACGCTGACAGATCTTGAAGAACTGGGGAATTCCTCGGTGAATCACCCTTTGACTAGT	2940
Qy	2941	GATGCAACCTGGTTCGTCTCTGATCTCGTTCCGATGGCTACATGTGGGAATGCCCC	3000
Db	2941	GATGCAACCTGGTTCGTCTCTGATCTCGTTCCGATGGCTACATGTGGGAATGCCCC	3000
Qy	3001	AACCTTTTAAGCTTCGGATGGAGAAACCTGGCAAGATCTCGACGGTCAATTTTCTGT	3060
Db	3001	AACCTTTTAAGCTTCGGATGGAGAAACCTGGCAAGATCTCGACGGTCAATTTTCTGT	3060
Qy	3061	CCACAGGATTTGACCCGATCCACGATGAGTTTCTCACTACGACAGCTCTGACAGATGC	3120
Db	3061	CCACAGGATTTGACCCGATCCACGATGAGTTTCTCACTACGACAGCTCTGACAGATGC	3120
Qy	3121	GGATATGTCGTGCAACAGCTTGAAGGAACGACCTTCCGCGTCTTGGCAGATTCAGCGAG	3180
Db	3121	GGATATGTCGTGCAACAGCTTGAAGGAACGACCTTCCGCGTCTTGGCAGATTCAGCGAG	3180
Qy	3181	CTGATTTCCGCGCATGAAATCTACGCAACGCGAGTTTCAGTAAACGGTTCGATGCTCGG	3240
Db	3181	CTGATTTCCGCGCATGAAATCTACGCAACGCGAGTTTCAGTAAACGGTTCGATGCTCGG	3240
Qy	3241	CTCGTGGGCTGGAATGGGGCTGCCCCGCCAGGATGATCACCCACAGTTGCAAGGAAGA	3300
Db	3241	CTCGTGGGCTGGAATGGGGCTGCCCCGCCAGGATGATCACCCACAGTTGCAAGGAAGA	3300
Qy	3301	TGGGAGCATGCTCTGACATGTCGCCCGGAGGTTATTTGGGCAACGACGGGATCTACAA	3360
Db	3301	TGGGAGCATGCTCTGACATGTCGCCCGGAGGTTATTTGGGCAACGACGGGATCTACAA	3360
Qy	3361	GAGCTCCTTCTCCAGAGGGGGAGTCGGGGGTAACTGATCTGATTAAGGTTCTGAACCT	3420
Db	3361	GAGCTCCTTCTCCAGAGGGGGAGTCGGGGGTAACTGATCTGATTAAGGTTCTGAACCT	3420
Qy	3421	GTCCGAGTGAACATCCGAGGCAATTTTCCCTCGAGTGGGATGCTGTCCTTTGCTGTG	3480
Db	3421	GTCCGAGTGAACATCCGAGGCAATTTTCCCTCGAGTGGGATGCTGTCCTTTGCTGTG	3480
Qy	3481	GATCGTATGATGATCGTCGCGTGTGAGGTAAACCTGCGCAATTAAGTATCGCGGAC	3540
Db	3481	GATCGTATGATGATCGTCGCGTGTGAGGTAAACCTGCGCAATTAAGTATCGCGGAC	3540
Qy	3541	GATTAATACGCCATTGATGATTACTGCAAGTATGAGACAGGTTTCATTCGCTTTCCGGC	3600
Db	3541	GATTAATACGCCATTGATGATTACTGCAAGTATGAGACAGGTTTCATTCGCTTTCCGGC	3600
Qy	3601	CTTCAAAAGTGAACCTATTGAGATTAAGTCATTAATAAGGCTTTTGTGGCGAATTTGT	3660
Db	3601	CTTCAAAAGTGAACCTATTGAGATTAAGTCATTAATAAGGCTTTTGTGGCGAATTTGT	3660
Qy	3661	ACAAATACCTTGGCAAAATCCCTTGATC 3687	
Db	3661	ACAAATACCTTGGCAAAATCCCTTGATC 3687	

RESULT 7		
BD165105		
LOCUS	1983 bp	DNA
DEFINITION		linear
Novel polynucleotide.		
ACCESSION		
BD165105		
PAT 17-JAN-2003		

```

VERSION      BD165105.1 GI:27870917
KEYWORDS     JP 2002191370-A/2904.
SOURCE       unidentified
ORGANISM     unidentified
REFERENCE    1 (bases 1 to 1983)
AUTHORS      Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
              Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
TITLE        Novel polynucleotide
JOURNAL      Patent: JP 2002191370-A 2904 09-JUL-2002;
              KYOWA HAKKO KOGYO CO LTD
              OS Corynebacterium glutamicum
              JP 2002191370-A/2904
COMMENT      PD 09-JUL-2002
              PF 15-DEC-2000 JP 2000405096
              PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
              PI KEIKO OCHIAI,
              PI HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
              OZAKI
              PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
              C12N1/15,
              PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
              PC C12P13/08,
              PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC
              G01N33/566,
              PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),
              PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
              PC C12N5/00,C12N15/00
              CC Novel polynucleotide
              FH Key
              FT source
              Location/Qualifiers
              1..1983
              /organism="Corynebacterium glutamicum"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"

FEATURES
    source             Location/Qualifiers
    Query Match       31.4%; Score 1874.2; DB 6; Length 1983;
    Best Local Similarity 96.6%; Pred. No. 0;
    Matches 1915; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
24
```

Df	1501	ACCATATCTCAGGACACTTTTACCGGTGAAGCTATTGCACCTGACAGCGTCACGGATGCC	1568
Oy	5339	ATGTTTGCCAGCGGAAAGCTTGGCTCAAGTGTGTGGATCTGTCACCACCAAGGGGACGCTG	5398
Df	1561	ATGTTTGCCAGCGGAAAGCTTGGCTCGGGCGTTGCAATGCTCCAACCAAGGGGACGTTA	1620
Oy	5339	GTTTTCACCAAGTAAGCGGAAAGATCGTGGTGGCTTCCCATCTGGTCAAGCTTTGGCACTC	5458
Df	1621	GTTTTCTCCGGTAGAGTGAAGAATTTGTGGGCAATTCACATCGGCAATGCTTTCCGACATT	1680
Oy	5459	CGCACATAAGGCTGAGGATGGTTCATGTGGATCTTGTGATGTCACATTTGGTTTGGACAC	5518
Df	1681	CGCACATAAGGCTGAGGATGGTTCATGTGGATCTTGTGATGTCACATTTGGTTTGGACACA	1740
Oy	5519	GTAACCTTCACCGGACCGCACTTTAAACCGCTGAAGAAAGCAGGGCGATGAAGTCAAACCA	5578
Df	1741	GTAACCTTCACCGGACCGCACTTTAAACCGCTGAAGAAAGCAGGGCGATGAAGTCAAACCA	1800
Oy	5579	GGGGAGCTGCTGTGTGAATTGATATTGATGCCATTAAAGGCTGCAAGTTATGAGTAAAC	5638
Df	1801	GGGGAGCTGCTGTGTGAATTGATATTGATGCCATTAAAGGCTGCAAGTTATGAGTAAAC	1860
Oy	5639	ACGCCGATTTGTTTTCGATTTACAAAGAAAACCGGACCGTGAACATTAACGTTTGGGCG	5699
Df	1861	ACGCCGATTTGTTTTCGATTTACAAAGAAAACCGGACCGTGAACATTAACGTTTGGGCG	1920
Oy	5699	GAAATTGAAGCGGGAGCCACCTGCTCAACGTCGCAAGAAAGAACCGGTGCACCAACA	5758
Df	1921	GAAATTGAAGCGGGAGCCACCTGCTCAACGTCGCAAGAAAGAACCGGTGCACCAACA	1980
Oy	5759	CCA 5761	
Df	1981	CCA 1983	
RESULT 8			
AX122988			
LOCUS	AX122988	1983 bp	linear PAT 11-MAY-2001
DEFINITION	Sequence 2904 from Patent EP1108790.		
ACCESSION	AX122988		
VERSION	AX122988.1 GI:14040476		
KEYWORDS			
SOURCE			
ORGANISM	Corynebacterium glutamicum		
	Corynebacterium glutamicum		
	Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterineae;		
REFERENCE	1		
AUTHORS	Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K., Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.		
TITLE	Novel polynucleotides		
JOURNAL	Patent: EP 1108790-A 2904 20-JUN-2001;		
	KYOWA HAKKO KOGYO CO., LTD. (JP)		
FEATURES			
source	Location/Qualifiers		
	1..1983		
	/organism="Corynebacterium glutamicum"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:1718"		
ORIGIN			
Query Match	31.4%; Score 1874.2; DB 6; Length 1983;		
Best Local Similarity	96.6%; Pred. No. 0;		
Matches 1915; Conservative	0; Mismatches 68; Indels 0; Gaps 0;		
Oy	3779	ATGACACATTAAGGACCTGCGGCAAGCAACCTCCTGGCGCGACATTTGGCGGGAAGACAACCTT	3838
Df	1	ATGACACATTAAGGACCTGCGGCAAGCAACCTCCTGGCGCGACATTTGGCGGGAAGACAACCTT	60
Oy	3839	GTCGCGCGCGCACACTGTGTCAACGCGTTTACGCTCTGTGCTCAAGACACCAAGAGATGTG	3898
Df	61	GTCGCGCGCGCACACTGTGTCAACGCGTTTACGCTCTGTGCTCAAGACACCAAGAGATGTG	120
Oy	3899	GATCGCCAAAGTCTGAGATGATGATCCAGATCTGAAAGCACCTTTGAAACTGGCGGCAATG	3958

Db	121	GATGCCCAAAAGTCGTGATGATGATTCACAGTCTGAAAGCACGTTTGAAACGGGTGATATG	180
Qy	3959	TTCCAGATCATCTGTCGGGCCAGGCGAGTGTGATCATGTTTTCAAAGAACTCGATGACGCA	4018
Db	181	TTCCAGATCATCTGTCGGGCCAGGCGAGTGTGATCATGTTTTCAAAGAACTCGATGACGCA	240
Qy	4019	ACCTTCAAAGACATGCGTGTGTCCACAGAGCAGCTCAAAAGTGTGTGTGGCTTACACGGCC	4078
Db	241	ACCTTCAAAGACATGCGTGTGTCCACAGAGCAGCTCAAAAGTGTGTGTGGCTTACACGGCC	300
Qy	4079	AACGTGTTACACCGGTGTGTAAGGTAATTTGGCGGACATTTTCGTCGCCGTATTCGAATC	4138
Db	301	AACGTGTTACACCGGTGTGTGTAAGGTAATTTGGCGGACATTTTCGTCGCCGTATTCGAATC	360
Qy	4139	TTGGTTGGTGGCGGTCTGCTCATGAGCTATCAACATGTGTGGTGGCAGAGATCTGTTC	4198
Db	361	TTGGTTGGTGGCGGTCTGCTCATGAGCTATCAACATGTGTGGTGGCAGAGATCTGTTC	420
Qy	4199	GCTCCGACATCATCTGTGTGAGATGTTCCCTCAGATCAGCGGTGTGTCTGAGATGATCAAC	4258
Db	421	GCTCCGACATCATCTGTGTGAGATGTTCCCTCAGATCAGCGGTGTGTCTGAGATGATCAAC	480
Qy	4259	CTGATGGCATCTGGCGCGTGTGCGGCTTCTGTCAGATGCAAGTGTGTGTTTACCGGACAAG	4318
Db	481	CTGATGGCATCTGGCGCGTGTGCGGCTTCTGTCAGATGCAAGTGTGTGTTTACCGGACAAG	540
Qy	4319	CGTTTCGGTGGCAATGAGATTCTGTGGGCGCCGGCATTTGGTATGCGAGTGTGTTCCCAAC	4378
Db	541	CGTTTCGGCGGCATGAGATTCTGTGGCGCGCGGTAATGGTATGGCGATGTGTTCGAGC	600
Qy	4379	CTGCTTAA CGGCTTACGAGTGTGGCGCCGACACATGACCGCGGGCGAAATGCCAATGTGTCC	4438
Db	601	TTGGTGAAACGGCTTACGAGTGTGGCGCCGACACATGAGCTCGGGCGAAATGCCAATGTGTCC	660
Qy	4439	CTGTTTGGTTGGATGATGTTGCTCAAGCTGTGTTTACAAGGCAACCGGTCTTCTGTGTGTTG	4498
Db	661	CTGTTTGGTTTGAAGTGTGCTTCCCAAGCCGGTTTACAAGGCAACCGGTCTTCTGTGTGTTG	720
Qy	4499	GTCCTTGGATTTCTGGCAACGATCGAGAAAGTTCTGCGACAAGCAGACTATGGGCACTGCA	4558
Db	721	GTTTCTTGGATTTCTGGCAACGATCGAGAAAGTTCTGCGACAAGCAGACTATGGGCACTGCA	780
Qy	4559	GACTTCTGATCACCCGAGTGTGACTCTGTGTCTACCGGGCTTCTTACGTTCAATGCT	4618
Db	781	GACTTCTGATCACCTCCAGTGTCTGACTGATGCTGTCTACCGGGATTCCTTACATTCATGCC	840
Qy	4619	ATTGTGTCAGAAATGCGTGTGGTGGGTGATCTTGTCTGACACCGGTCTGCAAGGACTCTAT	4678
Db	841	ATTGTGTCAGAAATGCGTGTGGTGGGTGATCTTGTCTGACACCGGTCTTACAGGCACTTAT	900
Qy	4679	GATTTCCGTTGCTCCAGTCCGCGGCTGTGCTTTCGATCTGAGCTTACCTCAACCAATGTTATC	4738
Db	901	GATTTCCGTTGCTCCAGTCCGCGGCTGTGCTTTCGATCTGAGCTTACCTCAACCAATGTTATC	960
Qy	4739	ACTGCTGTGCAACAGTCTCTTCCCGCCAAATTAGCTGAGACTGTTCAACCAAGGTGATTC	4798
Db	961	ACTGCTGTGCAACAGTCTCTTCCCGCCAAATTAGCTGAGACTGTTTAAACCAAGGTGATTC	1020
Qy	4799	TTCAATCTTCGAAACCGCATTCATGGCCAAATATGCGCGAGGGTGCAGCATGTTTGGCAGTG	4858
Db	1021	TTCAATCTTCGAAACCGCATTCATGGCCAAATATGCGCGAGGGTGCAGCATGTTTGGCAGTG	1080
Qy	4859	TTCTTCCAGGAAGAGTGAAGAACTCAAGGGCTTGCAGAGTGTCTTCAAGGTGTCTCCGCT	4918
Db	1081	TTCTTCCAGGAAGAGTGAAGAACTCAAGGGCTTGCAGAGTGTCTTCAAGGTGTCTCCGCT	1140
Qy	4919	GTTCTTGGTATTAAGAGCCTGTGCATCTTGTGTGAACTTTCGCTGTGCGCTGTGCGTTC	4978
Db	1141	GTTCTTGGTATTAAGAGCCTGTGCATCTTGTGTGAACTTTCGCTGTGCGCTGTGCGTTC	1200
Qy	4979	TACATTTGGTATCGGTAACCGCAGCTATGTGTGTGGCTTGAATGCACTCTTGAATATCAAG	5038
Db	1201	TTCAATCGGTAATCGGTAACCGCAGCTATGTGTGTGGCTTGAATGCACTCTTGAATATCAAG	1260

QY	5039	GCAGTTGCGTTGGGCGCTGGACAGTTTCTGTGGAGTTGTGTTCTATATGATGCTCCAGATATG	5098
Db	1261	GCATTTGCGTTGGGCGCTGGACAGTTTCTGTGGAGTTGTGTTCTATATGATGCTCCAGATATG	1320
QY	5099	GTCAATGTTCTTGGTGGGCGGTAGTACCTTTGTCATGCGATTTGGGCGACGATTTGCT	5158
Db	1321	GTCAATGTTCTTGGTGGGCGGTAGTACCTTTGTCATGCGATTTGGGCGACGATTTGCT	1380
QY	5159	TATGGCCCTTAACTTGTGTTCCGCCGAAGGCAAGCATTTGATCCAGATGCAACCGCTGCTCCA	5218
Db	1381	TATGGCCCTTAACTTGTGTTCCGCCGAAGGCAAGCATTTGATCCAGATGCAACCGCTGCTCCA	1440
QY	5219	GTGCTCTGAGGAACGACCAAAAGCCGAAGACAGAACACCCCGAGATTTTCAAAAGATTTCC	5278
Db	1441	GTGCTCTGAGGAACGACCAAAAGCCGAAGACAGAACACCCCGAGATTTTCAAAAGATTTCC	1500
QY	5279	ACCATCATCCAGGACACCTTTGACCGGTGAAGCTATGCACTGAGCAGCGTCAAGCATGCC	5338
Db	1501	ACCATCATCCAGGACACCTTTGACCGGTGAAGCTATGCACTGAGCAGCGTCAAGCATGCC	1560
QY	5339	ATGTTTTCGACAGCGAAAGCTTGGCTGCTCAAGTGTTCGATTCGTCGCCACCAAGGGGCACTG	5398
Db	1561	ATGTTTTCGACAGCGAAAGCTTGGCTGCTCAAGTGTTCGATTCGTCGCCACCAAGGGGCACTTA	1620
QY	5399	GTTTCACACAGTGAAGCGGAAAGATCGTGGTGGCTTCCCATCTGGTCAAGCTTTGCGAGTC	5458
Db	1621	GTTTCCTCCGGTGAATGGAAAGATTTGGTGGCATTTCCATCTGGCAGTGGCTTTGCGAGTT	1680
QY	5459	CGACCTAAAGCTGAGGATGTTTCCATATGTGATATCTTTGATGCACTTTGGTTTGCATCC	5518
Db	1681	CGACCTAAAGCTGAGGATGTTTCCATATGTGATATCTTTGATGCACTTTGGTTTGCATCA	1740
QY	5519	GTAACCTCCACAGCGACGCACTTTAAACCGCTGAAGAAAGCAGGGCGATGTAAGTCAAAACA	5578
Db	1741	GTAACCTCCACAGCGACGCACTTTAAACCGCTGAAGAAAGCAGGGCGATGTAAGTCAAAACA	1800
QY	5579	GGGAGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAGGCTGCAAGTTTATGAGTAAAC	5638
Db	1801	GGGAGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAGGCTGCAAGTTTATGAGTAAAC	1860
QY	5639	ACGCCGATTTGTTGTTTTCGATTAACAAGAAAACCGACCTGTAAACACTTACGTTTGGGC	5698
Db	1861	ACGCCGATTTGTTGTTTTCGATTAACAAGAAAACCGACCTGTAAACACTTACGTTTGGGC	1920
QY	5699	GAAATTGAAGGGGGAGCCCAACTGCTCAACGTCGCAAGAAAGAAAGAGCGGTCCAGCAACA	5758
Db	1921	GAAATTGAAGGGGGAGCCCAACTGCTCAACGTCGCAAGAAAGAAAGAGCGGTCCAGCAACA	1980
QY	5759	CCA 5761	
Db	1981	CCA 1983	
RESULT 9			
BD094213			
LOCUS	BD094213	1656 bp	DNA linear
DEFINITION	Genes for heat resistant enzymes of amino acid biosynthetic pathway		PAT 27-AUG-2002
ACCESSION	BD094213		
VERSION	BD094213.1	GI:22639801	
KEYWORDS	WO 0125447-A/71.		
SOURCE	Corynebacterium thermoaminogenes		
ORGANISM	Corynebacterium thermoaminogenes		
REFERENCE			
AUTHORS	Hirano,S., Nonaka,G., Matsuzaki,Y., Akiyoshi,N., Nakamura,K., Kimura,E., Osumi,T., Matsui,K., Kawahara,Y., Kurahashi,O., Nakamatsu,T. and Sugimoto,S.		
TITLE	Genes for heat resistant enzymes of amino acid biosynthetic pathway		
JOURNAL	Derived from thermophilic coryneform bacteria		
	Patent: WO 0125447-A 71 12-APR-2001;		

COMMENT	ORIGIN	FEATURES
<p> AJINOMOTO CO INC, SEIRO HIRANO, GEN NONAKA, YUMI MATSUZAKI, NAKIKO AKIYOSHI, KANAE NAKAMURA, EIICHIRO KIMURA, TSUYOSHI OSUMI, KAZUHIKO MATSUI, YOSHIO KAMAHARA, OSAMU KURAHASHI, TSUYOSHI NAKAMATSU, SHINICHI SUGIMOTO </p> <p> OS Corynebacterium thermoaminogenes PN WO 0125447-A/71 PD 12-APR-2001 PF 04-OCT-2000 WO 2000JP006913 PR 04-OCT-1999 JP 99P 282716, 01-NOV-1999 JP 99P 311147 PR PI SEIRO HIRANO, GEN NONAKA, YUMI MATSUZAKI, NAKIKO AKIYOSHI, KANAE NAKAMURA, PI EIICHIRO KIMURA, TSUYOSHI OSUMI, KAZUHIKO MATSUI, YOSHIO KAMAHARA, PI OSAMU KURAHASHI, TSUYOSHI NAKAMATSU, SHINICHI SUGIMOTO PC C12N15/60, C12N15/54, C12N15/53, C12N15/31, C12N15/56, C12N9/88, PC C12N9/12, PC C12N9/04, C07K14/34, C12N9/26, C12N13/04 </p>	<p> Query Match 24.4%; Score 1457; DB 6; Length 1656; Best Local Similarity 93.2%; Pred. No. 0; Matches 1545; Conservative 0; Mismatches 110; Indels 2; Gaps 2; </p>	<p> source Location/Qualifiers FT CDS Key Location/Qualifiers 1..1656 /organism="Corynebacterium thermoaminogenes" /mol_type="genomic DNA" /db_xref="taxon:161484" </p>
<p> 2031 ACGCTGCATTTCTTAAACACATCGAATGGAAGTCTCCACACCGCGTCAACCAGGTTTGG 2030 1 ACGGCCCATCTTTCACACATATCGAAGAGGCTCCACCGCGTTCACCGAGGCTTGG 60 2091 GCACCTTTGTCGGCGGCGCAAAACATCGTTGGTGGGCAACGTGGGAAGAAACCGCAC- 2149 61 GTACTTTGTCCCGCGGCGAATAATTCGTTGGTGGGCAACGTGGCAAGAAACAGACA 120 2150 GCCATCGCGGAACTGTGAAGGCGCGACGTGCTTCTTGCACAGGTTCCATCTGTAG 2209 121 GCCATCGCGGAACTGTGAAGGCGCGACGTGCTTCTTGCACAGGTTCCATCTGTCAA 180 2210 ATGCAACAATGCCACATCATCTGTGGATGAAGACAGCATGTTCACAGCTGGAAAACGCT 2265 181 ATGCAACAACGCGCACCATCATCTGTGGATGAAGACAGCATGTTCACAGCTGGAAAACGCT 240 2270 GATCACTACCGTTCATGAGAGCAATTAAGCTGGCTTAGAAACAAAAGAAAGTAATGT 2329 241 GACCATTAACCGTTCATGAGAGCAATTAAGCTGGCTTAGAAACAAAAGAAAGTAATGT 300 2330 GTGGGGCTATGCACACAGAACTTCCAGTTTCCGCTCGCTGATCATGTGACTCTCCGC 2389 301 GTGGGGCTATGCACACAGAACTTCCAGTTTCCGCTCGCTGATCATGTGACTCTCCGC 360 2390 AGGCGAGCTCATGATGCCAAGGAATGTCGTGATGAGATACCTTCCAGCTTAACT 2449 361 AGGCGAGCTCATGATGCCAAGGAATGTCGTGATGAGATACCTTCCAGCTTAACT 420 2450 ACCGACAGATCCAGAGTTTCCCTTGGACCAAGAGCGACCGGCTGGGCTCACACACA 2509 421 ACCGACAGATCCAGAGTTTCCCTTGGACCAAGAGCGACCGGCTGGGCTCACACACA 480 2510 CGCGCTTGAACCGGACCGGAGATTTGCAAGTGAAGCACTGCGCCGACGCTTTTACCGCG 2569 481 CGCGCTTGAACCGGACCGGAGATTTGCAAGTGAAGCACTGCGCCGACGCTTTTACCGCG 540 2570 ATGATCTTATGACCTGGATGATGCTATTCCGCTGAACCGGCTTTTACTAGCGCACAC 2629 541 ATGATCTTATGACCTGGATGATGCTATTCCGCTGAACCGGCTTTTACTAGCGCACAC 600 2630 TTTAACTTTTCAACCGGCAACTTAAATTTGACGAAAGCGCGCGCCACCAAAACC 2689 </p>		

Db 421 TGATTTGCGTGGTCCAGTCCGCGCTGCTGCTTCCGCTGCTGCTACTCACTCAATGCTCAT 480
 QY 4738 CACTGGTCTGCACCAAGTCTTCCCGCAATTGAGCTGGAGCTGTTCAACCAAGGTGGATC 4797
 Db 481 CACTGGTCTGCACCAAGTCTTCCCGCAATTGAGCTGGAGCTGTTCAACCAAGGTGGATC 540
 QY 4798 CTTCACTCTTCGCAACCGCATCCATGAGCCATATCCGCGAGGGGGGAGAGATGTTGGCACT 4857
 Db 541 CTTCACTCTTCGCAACCGCATCTATGAGCTATATGCCCGAGGGGTGGCGCATGTTGGCACT 600
 QY 4858 GTTCTTCTTCGCAACGAGTGAAGAGTCAAGGGCTTTCGAGGTGCTTCAGTGTCTCCG 4917
 Db 601 GTTCTTCTTCGCAACGAGTGAAGAGTCAAGGGCTTTCGAGGTGCTTCAGTGTCTCCG 660
 QY 4918 TGTTCCTGATATTAACAGAGCTTCGATCTTCGCTGTTGAACCTTCGCTTCGCTGCTGCT 4977
 Db 661 TGTTCCTGATATTAACAGAGCTTCGATCTTCGCTGTTGAACCTTCGCTTCGCTGCTGCT 720
 QY 4978 CTACATTTGATATTCGATCCGAGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5037
 Db 721 CTTCACTCGTATCCGATCCGAGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 5038 GGCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5097
 Db 781 GGCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 5098 GGTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5157
 Db 841 GGTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 5158 TTATGAGCTCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5217
 Db 901 TTATGAGCTCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 5218 AGTGCCTGAGGAAAGCAACCAAGCCGAGACAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 5277
 Db 961 AGTGCCTGAGGAAAGCAACCAAGCCGAGACAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1020
 QY 5278 CACCATCATCCAGGCACTTTTGAACCGGTGAAGCTATCCAGTGAAGCAAGCTGCAAGCTGCAAGCTG 5337
 Db 1021 CACCATCATCCAGGCACTTTTGAACCGGTGAAGCTATCCAGTGAAGCAAGCTGCAAGCTGCAAGCTG 1080
 QY 5338 CATGTTTTCGAGCGGAAAGCTTGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5397
 Db 1081 CATGTTTTCGAGCGGAAAGCTTGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 QY 5398 GGTTCACCAAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 5457
 Db 1141 AGTTCTCCGATGAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1200
 QY 5458 CCGCATTAAGCTGAGAGATGTTCCATGTGATCTTGAATGATGATGATGATGATGATGATGATGATGAT 5517
 Db 1201 TCGCACCAAGGCTGAGATGTTCCATGTGATCTTGAATGATGATGATGATGATGATGATGATGATGAT 1260
 QY 5518 CGTAAACCTCAACGAGCACTTTTAAACCGGCTGAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5577
 Db 1261 AGTAAACCTCAACGAGCACTTTTAAACCGGCTGAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
 QY 5578 AGGAGAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5637
 Db 1321 AGGAGAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 5638 CACGCGGATGTTGTTTGAATTAACAGAAACCGGACCTGTAACACTTACGTTTGGG 5697
 Db 1381 CACGCGGATGTTGTTTGAATTAACAGAAACCGGACCTGTAACACTTACGTTTGGG 1440
 QY 5698 CGAAATTAAGGAGGAGCAACCTGCTCAACGTCGAAGAAAGAAAGAGGAGGAGGAGGAGGAGGAGGAG 5757
 Db 1441 CGAAATTAAGGAGGAGCAACCTGCTCAACGTCGAAGAAAGAAAGAGGAGGAGGAGGAGGAGGAGGAG 1500
 QY 5758 ACCATTAAGTGAACCTTGAAGTGTTCG 5784
 Db 1501 ACCATTAAGTGAACCTTGAAGTGTTCG 1527

RESULT 12
 AX065221 1342 bp DNA linear PAT 24-JAN-2001
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1 Pompeius, M., Kroeger, B., Schroeder, H., Zelder, O. and Habermann, G.
 Corynebacterium glutamicum genes encoding proteins involved in
 carbon metabolism and energy production
 Patent: WO 010844-A 347 04-JAN-2001;
 BASF AKTIEGESELLSCHAFT (DE)
 JOURNAL
 FEATURES
 source
 CDS
 1..1342
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101..>1342
 /note="unassigned protein product; RXA00041"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC25850.1"
 /db_xref="GI:12542934"
 /translation="MHTLSLRPAVHYVTPPOGRINDPNMGVYVDGTLHVVYQHDGFP
 PPAKRTGMAHTTTPLEPQRLQWHLIDALYRPMDSYLDGYSAGAFPTDGLTFY
 TGNLKDGRKRAHONLVEDEPTGIMGSIHRSRPNPLIDGPASFTPHYRPMISPD
 GDGMWVIVGAQRENLTGAVALYRSTDLNWEPSGRTDLSDAQGSAADLVPGGYM
 BCPNLFITDEBETGEDLVLIJFCQGLRIHDEYTVHASSDQGVVGVKLBETGRVL
 RGFSELDGFHEFAPQVAVNGSDAVLVCMMGLPADDDHPTVAREGMVHCLTVRKLHL
 RNHAIVQELLPEGSGVIRSVLSEPRVDIRNLSLEMGVRSLSVRCGDRVAEV
 KPGLVITADMDTAIEITAGDQ"
 ORIGIN
 Query Match 22.1%; Score 1321.2; DB 6; Length 1342;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1329; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 2238 ATGAAGCAGCACTATCCAGCTGTAAGAAAGCTGATCACTACCGTTCATGAGCAATTA 2297
 Db 1 ATGAAGCAGCAGCATCCAAAGCTGTAAGAAAGCTGATCACTACCGTTCATGAGCAATTA 60
 QY 2298 AGCTGCGCTAGAAACAAAGAAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2357
 Db 61 AGCTGCGCTAGAAACAAAGAAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 QY 2358 TTTGGCGCT 2417
 Db 121 TTTGGCGCT 180
 QY 2418 GTACGTCGATGAGATCCCTTCACGCTTACTACACGACGATCCAGGTTTCCCTTGGC 2477
 Db 181 GTACGTCGATGAGATCCCTTCACGCTTACTACACGACGATCCAGGTTTCCCTTGGC 240
 QY 2478 ACCAAGGAGCAACCGGCTGGGCTCAACACACGCGCTGAGCCGAGACGAGCAATGCA 2537
 Db 241 ACCAAGGAGCAACCGGCTGGGCTCAACACACGCGCTGAGCCGAGACGAGCAATGCA 300
 QY 2538 GTGACGACACTGACCGAGCTCTTAAACCGGATGCATCTATGACTGATGATGCTA 2597
 Db 301 GTGACGACACTGACCGAGCTCTTAAACCGGATGCATCTATGACTGATGATGCTA 360
 QY 2598 TTCCGCTGAGCGCTATTTACTGACGACACTTAACTTTTCTTACACCGGCAACTTAA 2657
 Db 361 TTCCGCTGAGCGCTATTTACTGACGACACTTAACTTTTCTTACACCGGCAACTTAA 420
 QY 2658 AATTACGAGAAAGCGCGCGCAACCAAACTTGTGCAAGTCAGAGCAACCAACTGGGCT 2717

Db	482	TCAGCCCTGATAGTGGATGTTGGAAAAATGGTTCTTGGGGCCCAACGGAATAACTCAACCG	541
Qy	2867	GTGCACGCGGTTCTATACCGCTCGACAGATCTTGAATCTGGAAATTCCTCGGTGAATCA	2926
Db	542	GTGCACGCGGTTCTATACCGCTCGACAGATCTTGAATCTGGAAATTCCTCGGTGAATCA	601
Qy	2927	CCTTGACCTCAGGATGACACAACCTGGTTCGGCTCCGATCTGGTCCGATGCGATCA	2986
Db	602	CCTTGACCTCAGGATGACACAACCTGGTTCGGCTCCGATCTGGTCCGATGCGATCA	661
Qy	2987	TGTGGGAATGCCCCAACCCTTTTACGCTTCGCGATGAAGAACTGGCAATCTCGACG	3046
Db	662	TGTGGGAATGCCCCAACCCTTTTACGCTTCGCGATGAAGAACTGGCAATCTCGACG	721
Qy	3047	TGCTGATTTTCTGTCACAAAGATTTGACCGGAATTCACGATGAGGTTACTCACTACGAA	3106
Db	722	TGCTGATTTTCTGTCACAAAGATTTGACCGGAATTCACGATGAGGTTACTCACTACGAA	781
Qy	3107	GCTCTGACCAATGCGGGATATGTGCTCGACAAAGCTTGAAAGAACCACTTCGCGCTCTTCG	3166
Db	782	GCTCTGACCAATGCGGGATATGTGCTCGACAAAGCTTGAAAGAACCACTTCGCGCTCTTCG	841
Qy	3167	GAGGATTCAGGAGAGCTGGAATTCGCGCCATGAATCTACGACCCGAGGTTGACAGAAACG	3226
Db	842	GAGGATTCAGGAGAGCTGGAATTCGCGCCATGAATCTACGACCCGAGGTTGACAGAAACG	901
Qy	3227	GTTCTGATGCTCGGTGCTCGTGGGCTTGATGAGGGCTTCGCCGACAGATGATCAACCAACG	3286
Db	902	GTTCTGATGCTCGGTGCTCGTGGGCTTGATGAGGGCTTCGCCGACAGATGATCAACCAACG	961
Qy	3287	TTGCACAGGAAGATGGGTGCACTGCTGACTGTGCCCCGCAAGCTTCATTTGGCAAC	3346
Db	962	TTGCACAGGAAGATGGGTGCACTGCTGACTGTGCCCCGCAAGCTTCATTTGGCAAC	1021
Qy	3347	ACGGGATCTACCAAGAGCTCCTTCCCGAAGGGGGAATGGGGGTAATCAGATCTGAT	3406
Db	1022	ACGGGATCTACCAAGAGCTCCTTCCCGAAGGGGGAATGGGGGTAATCAGATCTGAT	1081
Qy	3407	TAGGTTCTGAACTGTCCGAGTAGACATCCGAGGCATATTTCCCTCGAGTGGATGGTG	3466
Db	1082	TAGGTTCTGAACTGTCCGAGTAGACATCCGAGGCATATTTCCCTCGAGTGGATGGTG	1141
Qy	3467	TCCGTTTGTCTGTGATGCTGATAGGTGATCGTCGCTAGTACTGAGGTAAACCTGCGCAAT	3526
Db	1142	TCCGTTTGTCTGTGATGCTGATAGGTGATCGTCGCTAGTACTGAGGTAAACCTGCGCAAT	1201
Qy	3527	TAGGATTCGCGGACGATTAATACGCCATTTGAGTAACTGCGAGGTGATGGAACAAGTTTCAT	3586
Db	1202	TAGGATTCGCGGACGATTAATACGCCATTTGAGTAACTGCGAGGTGATGGAACAAGTTTCAT	1261
Qy	3587	TCCGTTTCCGGGCTTCAAGAGTGACACTATTGAGACA	3625
Db	1262	TCCG-TTTCCGGGCTTCAAGAGTGACACTATTGAGACA	1299
RESULT 14			
AX122989			
LOCUS			
DEFINITION Sequence 2905 from Patent EP1108790.			
VERSION AX122989			
KEYWORDS AX122989.1 GI:14040477			
SOURCE			
ORGANISM			
Corynebacterium glutamicum			
Corynebacterium glutamicum			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
REFERENCE			
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,			
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.			
Novel polynucleotides			
Patent: EP 1108790-A 2905 20-JUN-2001;			
KIOWA HAKKO KOGYO CO., LTD. (JP)			
TITLE			
JOURNAL			

FEATURES		Location/Qualifiers
SOURCE		1..1299
		/organism="Corynebacterium glutamicum"
		/mol_type="unassigned DNA"
		/db_xref="taxon:1718"
ORIGIN		
Query Match	21.2%;	Score 1267.8; DB 6; Length 1299;
Match Local Similarity	99.0%;	Pred. No. 0;
Matches 1286;	Conservative 0;	Mismatches 12; Indels 1; Gaps 1;
QY	2327	TGTTGTGGGGCTATGTSCACACAGAACTTTCAGATTGCGCCCTGCGTACCATATGTGACTCTTC 2386
DB	2	TGTGTGGGGCTATGTSCACACAGAACTTTCAGATTGCGCCCTGCGTACCATATGTGACTCTTC 61
QY	2387	CGCAGGGCAGGCTCAATGATCCCAACGGAAATGTAAGTCGATGAGATACCCTCCAGCTCT 2446
DB	62	CGCAGGGCAGGCTCAATGATCCCAACGGAAATGTAAGTCGATGAGATACCCTCCAGCTCT 121
QY	2447	ACTTACCGACGACGATCCGAGTTTCCCTTTCGACCAAAAGCGCACCGGCTGGGCTTACACCA 2506
DB	122	ACTTACCGACGACGATCCGAGTTTCCCTTTCGACCAAAAGCGCACCGGCTGGGCTTACACCA 181
QY	2507	CCACGCGGTTGACCGGACCGGACGATTTGACAGTGAAGCAGACCTGCGCGACGCTCTTTACC 2566
DB	182	CCACGCGGTTGACCGGACCGGACGATTTGACAGTGAAGCAGACCTGCGCGACGCTCTTTACC 241
QY	2567	CGGATGCACTCTATGACCTGTGATGATGCTATTCCGCTGAGACCGTATTTTACTGACGGCA 2626
DB	242	CGGATGCACTCTATGACCTGTGATGATGCTATTCCGCTGAGACCGTATTTTACTGACGGCA 301
QY	2627	CACTTAAACTTTTCTACACCGGCAACCTTAAATAATGACGGAAGCGCCGCGCACCCCAA 2686
DB	302	CACTTAAACTTTTCTACACCGGCAACCTTAAATAATGACGGAAGCGCCGCGCACCCCAA 361
QY	2687	ACCTGTTCGAAGTGTGAGAGACCCAACTGAGGCTGATGGGCGGCAATTATCGCGCTGGGCTTA 2746
DB	362	ACCTGTTCGAAGTGTGAGAGACCCAACTGAGGCTGATGGGCGGCAATTATCGCGCTGGGCTTA 421
QY	2747	AAATCCGCTTATCGACGACCCGCGACGCGTTTCAACCCATTACCGGATCCCATGA 2806
DB	422	AAATCCGCTTATCGACGACCCGCGACGCGTTTCAACCCATTACCGGATCCCATGA 481
QY	2807	TCAGCCCTGATGATGATGTTGTAAGAACATGATCTTTGGGGCCCAACGCGAAACCTCACCG 2866
DB	482	TCAGCCCTGATGATGATGTTGTAAGAACATGATCTTTGGGGCCCAACGCGAAACCTCACCG 541
QY	2867	GTGACGCGGTTCTTATACCGCTCCACAGATCTTGAAAACTGGGAATTTCTCCGTTGAAATCA 2926
DB	542	GTGACGCGGTTCTTATACCGCTCCACAGATCTTGAAAACTGGGAATTTCTCCGTTGAAATCA 601
QY	2927	CCTTTGAACCTCAGTATGACCAACCTGGTCTGCTGCTGATCTTCGTTCCGATGGCTTACA 2986
DB	602	CCTTTGAACCTCAGTATGACCAACCTGGTCTGCTGCTGATCTTCGTTCCGATGGCTTACA 661
QY	2987	TGTGGGAATGCCCCCAACCTTTTACGCTTCCGATGGAAGAACTGGGAAGATCTCGACG 3046
DB	662	TGTGGGAATGCCCCCAACCTTTTACGCTTCCGATGGAAGAACTGGGAAGATCTCGACG 721
QY	3047	TGCTGATTTTCTGTCCCAAGATTTGACCGAATCCAGATGAGGTTATCTCATACGCCAA 3106
DB	722	TGCTGATTTTCTGTCCCAAGATTTGACCGAATCCAGATGAGGTTATCTCATACGCCAA 781
QY	3107	GCCTGACCAAGTGGGATATGTGTGACAAAGCTGGAAGGAAGCACTTCGCGGCTTTGCG 3166
DB	782	GCCTGACCAAGTGGGATATGTGTGACCAAGCTTGAAGGAAGCACTTCGCGGCTTTGCG 841
QY	3167	GAGGATTCACGACGCTGATTTTGGCCATGAATTTCTACGACCGCAGGTTGCACTAAACG 3226
DB	842	GAGGATTCACGACGCTGATTTTGGCCATGAATTTCTACGACCGCAGGTTGCACTAAACG 901
QY	3227	GTTCGTATGCTGGCTGTGTGGCTGATGGGCGCTGCCCGCAGATGATCACCCAACAG 3286
DB		

Df		902	GTTCTGATGCCTGGCTGTGGAGCTGGANATGGSGAGCTGCCCGCAGAGATGATCAACCAACAG	961
OY		3287	TTTGACAAGANGATGGGTGCACTGGCTGACTGTGGCCCCCGAAGCTTCATTTGGGCAACC	3346
Df		962	TTTGACACGGGAAGATGGGTGTGCACTGGCTGACTGTGGCCCCCGAAGCTTCATTTGGGCAACC	1021
OY		3347	AACGATCTACCAAGAGCTCCCTCTCCAGAGGGGGAGTCCGGGGGTATCATGCTCTGAT	3406
Df		1022	ACGGATCTTACAGAGCTTCTTCTCCAGAGGGGGAGTCAAGGGGTATCATGCTCTGAT	1081
OY		3407	TAGGTTCTGAACCTGTCCGAGTAGACATCCGAGGCAATATTTCCCTCGAGTGGAGTGTG	3466
Df		1082	TAGGTTCTGAACCTGTCCGAGTAGACATCCGAGGCAATATTTCCCTCGAGTGGAGTGTG	1141
OY		3467	TCCGTTTCTGTGGATGTGTGATGTGTATGCTGTGCTGAGTGAAGATAAACTTGGCGAAT	3526
Df		1142	TCCGTTTCTGTGGATGTGTGATGTGTATGCTGTGCTGAGTGAAGATAAACTTGGCGAAT	1201
OY		3527	TAGTATCGCGGACCATTAATACAGCCATTGAGTAACTGACAGTGTATGACAGGTTTCAT	3586
Df		1202	TAGTATCGCGGACCATTAATACAGCCATTGAGTAACTGACAGGTTATGACAGGTTTCAT	1261
OY		3587	TCCGTTTTCCGGGCTTCAAAGGTGACACTATTGAGAGA	3625
Df		1262	TCCG-TTTCGGGCTTCAAAGGTGACACTATTGAGAGA	1299
RESULT 15				
LOCUS	AX065287	1287 bp	DNA	linear PAT 24-JAN-2001
DEFINITION	Sequence 413 from Patent WO0100844.			
ACCESSION	AX065287			
VERSION	AX065287.1 GI:12542999			
KEYWORDS	.			
SOURCE	Corynebacterium glutamicum			
ORGANISM	Corynebacterium glutamicum Bacteria; Actinobacteriae; Actinobacteriales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
REFERENCE	1			
AUTHORS	Pompejus,M., Kroeger,B., Schroeder,H., Zelder,O. and Haberbauer,G.			
TITLE	corynebacterium glutamicum genes encoding proteins involved in carbon metabolism and energy production			
JOURNAL	Patent: WO 0100844-A 413 04-JAN-2001;			
FEATURES	BASF AKTIENGESellschaft (DB)			
	Location/Qualifiers			
	1..1287			
CDs	/organism="Corynebacterium glutamicum"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:1718"			
	101..1267			
	/note="unnamed protein product; KXN00043"			
ORIGIN	/codon_start=1			
	/transl_table=11			
	/protein_id="CAC25883.1"			
	/db_xref="GI:12543000"			
	/translation="MAEYVHOENAGAVKKIEGRIVPHGVINGFIQLENGIITLST GEPAKRNAGFHELPITYPSFDILHNKGANGAPRTGQDARBAQAHRERHGTVMY ASWABAPDALAQVEELIPCEBEGLGHIHQEPINACKGNPFRPGNPPTDI AQVIHAGKMWIKSIIVAREPDTNLTELIDLCAAHNIASFGHTADFPDTTALIAKE KNVTATATHLENAMPPLHHRDGSGVALIAARADAVEELIADGHVHADDTVDLAR NNAFITPDAMEAAMPDESYLIAGLVNTVTGVAVALRDGSAIAGSTSLAQFHNVAV RMTULIDATIHSTVAAKILIGLSDEHIEKSNPVNFVPDSNQQLQQVHLDHQV1"			
Query Match				
Best Local Similarity 97.6%; Pzed. No. 9.9e-307;				
Matches 1256; Conservative 0; Mismatches 31; Indels 0; Gaps 0;				
OY	242 AACAGAGGCGCCAGTCCGAGAGTAATTAACCTGATCCGAGACATTAAGACATCAATAC	301		
Df	1 AACHGAGGCGCTCAAGTCCGAGAGTAATTAACCTGATCCGAGACATTAAGACATCAATAC	60		
OY	302 GTTCATGCTTCTCGAAGAACCAAAATTAACCTCGAAGATGAGCAGAACTGTGCATTA	361		

Db	61	GTCTTAAGCTTGTCTGGAAAGAGCAATTAACCTCAGAAAGATGGCAGAAAGTGATGCACTTA	120
Qy	362	TCAGAAAAATGCAGGTCAAGCATTTAAAAAATTTAGGGAAGAATTGTTCCCCCTTCGG	421
Db	121	TCAGAAAAATGCAGGTCAAGCATTTAAAAAATTTAGAAAGAAATTTGTTACCCGCCACGG	180
Qy	422	GGTATATGAATGGCTTTCTCCAACTCGAAAAACGGGATCATACGGAACTCTCTGGAAAC	481
Db	181	GGTATATGAATGGCTTTCTCCAACTCGAAAAACGATCATACGGAACTCTCTGGAAAC	240
Qy	482	AGCATTAATAAACGCAAGATTCAACCCCGAATCTCCACGATTTGTTCCCGTTTATTGA	541
Db	241	AGCATTAATAAACGCAAGATTCAACCCCGAATCTCCACGATTTGTTCCCGTTTATTGA	300
Qy	542	TCTTCATTAATACGGTGGAAAACGGTGGCGGTTTCCTAACGGGAAACGAGAGCCAGGGAG	601
Db	301	TCTTCATTAATACGGTGGAAAACGGTGGCGGTTTCCTAACGGGAAACGAGAGCCAGGGAG	360
Qy	602	GAACACCGCGCAGTATCAACCGCAACATATGACACGCGTATGTTTGCACAGATGATTTC	661
Db	361	GAATGCGCGGAGATATCAACCGGAAACATATGACACGCGTATGTTTGCAGCATGATTTC	420
Qy	662	GACGCGCGCTGACGCACTGGCAGCGCAGAGTGGAAAACCTTATTTCCCTTGTGTGAAGAGT	721
Db	421	GACGCGCGCTGACGCACTGGCAGCGCAGAGTGGAAAACCTTATTTCCCTTGTGTGAAGAGG	480
Qy	722	CTGTGTGCGGCAATTCACCTCGAAGGCGCTTTCATCAACGATACGGTGTGTGTCTCA	781
Db	481	CTGTGTGCGGCAATTCACCTCGAAGGCTTTCATCAACGATACGGTGTGTGTCTCA	540
Qy	782	AAACCGGATTTCAATTTTCCCGCAACCAACGATTTGCGCGGTGATCCATGCGGG	841
Db	541	AAACCGGATTTTATTTTCCCGCAACCAACGATTTGCGCGGTGATCCATGCGGG	600
Qy	842	AAAAGGTGGATCAAAATGCATCACAAGTAGGCGCGGAAACCTGACATCTTTCTAGACTCT	901
Db	601	AAAAGGTGGATCAAAATGCATCACAAGTAGGCGCGGAAACCTGACATCTTTCTAGACTCT	660
Qy	902	CGATCTCTGCGCAGCGCAACATCATATTGCTCTTCTGGGCAACATGATGCAATTTTGA	961
Db	661	CGATCTCTGCGCAGCGCAACATCATATTGCTCTTCTGGGCAACATGATGCAATTTTGA	720
Qy	962	TACCATTAACAGCGCAATTTGCTTGTGCTTAAGAAAAATGTGACGGTCAAGGCTAAGCA	1021
Db	721	TACCATTAACAGCGCAATTTGCTTGTGCTTAAGAAAAATGTGACGGTCAAGGCTAAGCA	780
Qy	1022	TTTGTTCATGCGATGCTCTCGCTGCAATCATAGGAGCTCCCGGACGCGTGGCGCTTGGCT	1081
Db	781	TTTGTTCATGCGATGCTCTCGCTGCAATCATAGGAGCTCCCGGACGCGTGGCGCTTGGCT	840
Qy	1082	TGCTGCGGACGCTGCGGGGAACGATATGTTGAATGATGCGCCGACGGCGTGCATTTGGC	1141
Db	841	TGCTGCGGACGCTGCGGGGAACGATATGTTGAATGATGCGCCGCGTGCATTTGGC	900
Qy	1142	CGATGAAAACGGTCAATCTAGCTCGTTCGTAACAACGCTTTTTCATCAGGACGGCATGSA	1201
Db	901	CGATGAAAACGGTCAATCTAGCTCGTTCGTAACAACGCTTTTTCATCAGGACGGCATGSA	960
Qy	1202	AGCGCGCGGAATGCCAGACGCTGAGTACATTTTGGGCTTTTGAACGTACCCGTCAACGA	1261
Db	961	AGCGCGCGGAATGCCAGACGCTGAGTACATTTTGGGCTTTTGAACGTACCCGTCAACGA	1020
Qy	1262	TGGATGCGCCGCTGCGCGGATATGGGGGCGCATGCGCGGGGGCAACACATCATAGCGAG	1321
Db	1021	TGGATGCGCCGCTGCGCGGATATGGGGGCGCATGCGCGGGGGTACACACATAGCGAG	1080
Qy	1322	TCAGTTCTGTGACACAGTGGCGCAGGGGATATGACGTTTATCGACGGGACCTTCCACACTC	1381
Db	1081	TCAGTTCTGTGACACAGTGGCGCAGGGGATATGACGTTTATCGACGGGACCTTCCACACTC	1140
Qy	1382	AACGCTGCGCGTAAAAATTTCTGCTTTGGGATATAGAAATCGCTAATCCAAACCTTGC	1441

Db	1141 AACCGTCGCCGCCAAATTCGGA	CTAGCATCAGAAATCGTTAAATCCAA	CCCTGT	1200
Qy	1442	AAATTTTGGTCTTTGACTCAAA	CGCCAGGTGCAAAAGTTCATTAG	GTCACT 1501
Db	1201	AAATTTTGGTCTTTGACTCAAA	CGCCAGGTGCAAAAGTTCATTAG	GTCACT 1260
Qy	1502	ACTTAAGTACGAGTAAACTAT	CTCTG	1528
Db	1261	AAATTTAAATACGAGCAAA	CTTCTCTG	1287

Search completed: March 10, 2005, 03:48:13
 Job time : 24435 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 13:35:43 ; Search time 2740 Seconds
(without alignments)
12895.962 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969

Sequence: 1 agtcgcgtgcagccaccatc.....tgatatcttcgcccgtgaa 5969

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: _geneseq19808:*
- 2: _geneseq19908:*
- 3: _geneseq20008:*
- 4: _geneseq20018a:*
- 5: _geneseq20018b:*
- 6: _geneseq20028a:*
- 7: _geneseq20028b:*
- 8: _geneseq20038a:*
- 9: _geneseq20038b:*
- 10: _geneseq20038c:*
- 11: _geneseq20038d:*
- 12: _geneseq20048a:*
- 13: _geneseq20048b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5969	100.0	5969	4	AAf32543 Brevibact
2	5656.2	94.8	349980	5	AAH68533 C glutam
3	3687	61.8	6911	2	AAf34177 Corynefor
4	1874.2	31.4	1983	5	AAH67869 C glutam
5	1457	24.4	1656	5	AAf87497 Coryneb
6	1412.6	23.7	1527	5	AAf31528 C glutam
7	1412.6	23.7	1527	6	AB655346 DNA encod
8	1321.2	22.1	1342	4	AAf71533 Coryneb
9	1267.8	21.2	1299	5	AAH67870 C glutam
10	1237.4	20.7	1287	4	AAf71567 Coryneb
11	1237.4	20.7	1287	4	AAf71567 Coryneb
12	1110.4	18.6	1152	5	AAH67872 C glutam
13	1049.8	17.6	1109	5	AAf31529 C glutam
14	762	12.8	882	4	AAf71571 Coryneb
15	671.8	11.3	759	5	AAH67871 C glutam
16	474.2	7.9	479	5	AAf87433 Coryneb
17	466.6	7.8	498	5	AAf87432 Coryneb
18	322.8	5.4	110000	10	ADP77343_05
19	294	4.9	1543	10	ADD13201 C. glutam
20	232.2	3.9	2913	4	AAH54100 S. epider

C 21	215.2	3.6	1059	4	AAf71769 Coryneb
C 22	215.2	3.6	1059	4	AAH56088 C. glutam
C 23	208	3.5	777	8	ACA30091 Prokaryot
C 24	192.6	3.2	1212	8	ACA30089 Prokaryot
C 25	192	3.2	975	6	ABK73538 Bacillus
C 26	183.2	3.1	1452	11	ABD00072 Klebsiell
C 27	181.4	3.0	1386	11	ABD00428 Klebsiell
C 28	181	3.0	1491	10	ADH84532 Enterococ
C 29	177.6	3.0	1434	8	ACF73163 Staphyloc
C 30	177.6	3.0	30246	2	AAV74367 Staphyloc
C 31	176.2	3.0	861	8	ACF74550 Staphyloc
C 32	174.2	2.9	1380	6	ABK73394 Bacillus
C 33	173.8	2.9	2127	8	ABK42306 Streptoco
C 34	173.8	2.9	110000	10	ABH556454_16
C 35	172.2	2.9	1965	10	ABX07668 S. pneumo
C 36	172.2	2.9	3695	2	AAV52334 Streptoco
C 37	168.8	2.8	906	13	ADS02113 Staphyloc
C 38	167.8	2.8	1932	13	ADR92012 Novel S.
C 39	163.6	2.7	1884	4	AAH55732 Streptoco
C 40	162.2	2.7	1428	11	ACH96011 Klebsiell
C 41	161	2.7	1380	5	AAH67868 C. glutam
C 42	157	2.6	11330	4	AAH59544 Propionib
C 43	157	2.6	11330	8	ACF64473 Propionib
C 44	154.2	2.6	9769	2	AAV52163 Streptoco
C 45	154.2	2.6	110000	10	ABH556454_15

ALIGNMENTS

RESULT 1	AAf32543	standard; DNA; 5969 BP.
ID	AAf32543	
XX	AAf32543;	
AC	XX	
DT	11-SEP-2003 (revised)	
DT	20-APR-2001 (first entry)	
DE	Brevibacterium lactofermentum sucrose PTS enzyme II DNA SEQ ID NO:1.	
XX	XX	
KW	Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;	
KW	phosphoenolpyruvate:carboxylate phosphotransferase system; glucose;	
KW	coryneform bacterium; phosphoenolpyruvate-sugar transport system; db.	
XX	XX	
OS	Corynebacterium glutamicum.	
XX	XX	
FH	Key	Location/Qualifiers
FT	CDS	3779..5764
FT		/*tag= a
FT		/product= "sucrose PTS enzyme II"
PN	MO200102584-A1.	
PD	11-JAN-2001.	
XX	XX	
PF	30-JUN-2000; 2000MO-JP004348.	
XX	XX	
PR	02-JUL-1999; 99JP-00189512.	
XX	XX	
PA	(AJIN) AJINOMOTO CO INC.	
XX	XX	
PI	Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;	
XX	XX	
DR	WPI: 2001-138150/14.	
XX	XX	
PT	P-PSDB; AAB69080.	
XX	XX	
PT	Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme	
PT	II obtained by cassette ligation-mediated amplification of downstream	
PT	domain of coryneform bacterium sucrose gene, with sucrose-binding	
XX	XX	
PS	Claim 3; Page 22-29; 45pp; Japanese.	

1921 TTCAAACTCTGGCATCGCGGAAACGGCAATCGCTTTCATTGAACCATCTCTCTG 1980
1921 TTCAAACTCTGGCATCGCGGAAACGGCAATCGCTTTCATTGAACCATCTCTCTG 1980
1981 GTGAGACTGACAAAGGTCCAGGCGCTGCACTTAAACTGTGAGAGCAACCGCTGAT 2040
1981 GTGAGACTGACAAAGGTCCAGGCGCTGCACTTAAACTGTGAGAGCAACCGCTGAT 2040
2041 CTTGAAACACCATGAAAGAGTCCCAACCGCGCTCAACCGAGGTTTGAGCACTTTGTC 2100
2041 CTTGAAACACCATGAAAGAGTCCCAACCGCGCTCAACCGAGGTTTGAGCACTTTGTC 2100
2101 CCGCGGCAAAACATGCTGTGGTGGCACTGTGTAAGAAAGCGGAGCCATCCGCGG 2160
2101 CCGCGGCAAAACATGCTGTGGTGGCACTGTGTAAGAAAGCGGAGCCATCCGCGG 2160
2161 AACTGTGAAGGCGCCAGTACTGCTTCTTGCCAGGTTCCATCTGTAGATGACAAACAT 2220
2161 AACTGTGAAGGCGCCAGTACTGCTTCTTGCCAGGTTCCATCTGTAGATGACAAACAT 2220
2221 GCCACATCATCTGTGATGAGAGCAGCATATCCAAAGTGAAGAAACGCTGATCACTACG 2280
2221 GCCACATCATCTGTGATGAGAGCAGCATATCCAAAGTGAAGAAACGCTGATCACTACG 2280
2281 TCTCATGAGCAATTAAAGCTGCGCTAGAAACAAAGAAAGTACTGTGTGGGCTTAG 2340
2281 TCTCATGAGCAATTAAAGCTGCGCTAGAAACAAAGAAAGTACTGTGTGGGCTTAG 2340
2341 CACACAGAACTTTCAGATTGCGCCCTGCGTACATGTGACTCTCCGAGAGGACGCTC 2400
2341 CACACAGAACTTTCAGATTGCGCCCTGCGTACATGTGACTCTCCGAGAGGACGCTC 2400
2401 AATGATCCAAACGGAATGTAAGTGAATGAGATACCTTCCAGCTTACTACAGACGAT 2460
2401 AATGATCCAAACGGAATGTAAGTGAATGAGATACCTTCCAGCTTACTACAGACGAT 2460
2461 CAGAGTTTCCCTTGGCAACAAAGCGCACCGGCTGGGCTCACACACACGCGCTTAGC 2520
2461 CAGAGTTTCCCTTGGCAACAAAGCGCACCGGCTGGGCTCACACACACGCGCTTAGC 2520
2521 GGAACCGACAGATTGAGTGAAGCAGCCTGCCAGCGCTCTTACCCGAGTGCATCTAT 2580
2521 GGAACCGACAGATTGAGTGAAGCAGCCTGCCAGCGCTCTTACCCGAGTGCATCTAT 2580
2581 GACTGAGATGATGCTATTCCGGTGAAGCGGATTTTACTGACGACACACTTAACTTTTC 2640
2581 GACTGAGATGATGCTATTCCGGTGAAGCGGATTTTACTGACGACACACTTAACTTTTC 2640
2641 TACACCGGCAACCTTAAATTTGACGAAAGCGCGGCCAACCTTGTGCAAGTC 2700
2641 TACACCGGCAACCTTAAATTTGACGAAAGCGCGGCCAACCTTGTGCAAGTC 2700
2701 GAGAACCCAACTGGGCTGATGGGCGGCACTTCACTGCGCTTGGCTTAAATCCGCTTATC 2760
2701 GAGAACCCAACTGGGCTGATGGGCGGCACTTCACTGCGCTTGGCTTAAATCCGCTTATC 2760
2761 GACGAGCCGCGCAGCGGTTTCAACCCCACTTACCGGATCCCATGATCAGCCCTGATGT 2820
2761 GACGAGCCGCGCAGCGGTTTCAACCCCACTTACCGGATCCCATGATCAGCCCTGATGT 2820
2821 GATGTGGAACAATGTTCTTGGGGGCCAAACGGAACCTCAACCGGTGACAGGTTCTA 2880
2821 GATGTGGAACAATGTTCTTGGGGGCCAAACGGAACCTCAACCGGTGACAGGTTCTA 2880
2881 TACCGCTGACAGATTTGAAAACCTGGAAATCTCCGCTGAATCACTTTGACCTCAAT 2940
2881 TACCGCTGACAGATTTGAAAACCTGGAAATCTCCGCTGAATCACTTTGACCTCAAT 2940
2941 GATGCAACAACCTGTGCTCTCTCTGATCTGTTCCGATGGCTACATGTGGGAATGCCCC 3000
2941 GATGCAACAACCTGTGCTCTCTCTGATCTGTTCCGATGGCTACATGTGGGAATGCCCC 3000

3001 AACCTTTTACGCTTCGAGTGAAGAACTGGGGAAGATCTCGACGTCGTGATTTCTGT 3060
3001 AACCTTTTACGCTTCGAGTGAAGAACTGGGGAAGATCTCGACGTCGTGATTTCTGT 3060
3061 CCACAGAGATTGACCGAATCCACAGATGAGTTACTCACTACGCAAGCTCTGACAGTGC 3120
3061 CCACAGAGATTGACCGAATCCACAGATGAGTTACTCACTACGCAAGCTCTGACAGTGC 3120
3121 GGAATGTGTGACAAAGCTTGAAGAACGACCTTCGGCTTCTTGAGAGATTCAGCGAG 3180
3121 GGAATGTGTGACAAAGCTTGAAGAACGACCTTCGGCTTCTTGAGAGATTCAGCGAG 3180
3181 CTGGAATTTGGGCAATTAATTCAGCACCGGAGTTGAGTAAACGTTCTGATGCTGAG 3240
3181 CTGGAATTTGGGCAATTAATTCAGCACCGGAGTTGAGTAAACGTTCTGATGCTGAG 3240
3241 CTTGTGGGCTGATGGGCTGCGCGGAGATGATCAACCAAGTTGACAGAGAGA 3300
3241 CTTGTGGGCTGATGGGCTGCGCGGAGATGATCAACCAAGTTGACAGAGAGA 3300
3301 TGGGTGCACTGCTGACTGTGCTCCGCAAGCTTCAATTTGCGCAACACGCGATCTAAC 3360
3301 TGGGTGCACTGCTGACTGTGCTCCGCAAGCTTCAATTTGCGCAACACGCGATCTAAC 3360
3361 GAGCTCCTTCTCCAGAGGAGGAGTCCGGGGAATCAGATCTGTAATGATGCTGAACCT 3420
3361 GAGCTCCTTCTCCAGAGGAGGAGTCCGGGGAATCAGATCTGTAATGATGCTGAACCT 3420
3421 GTCCAGTGAATCCAGGCAATATTTTCCCTGAGTGGAGTGTCCGTTGCTGTG 3480
3421 GTCCAGTGAATCCAGGCAATATTTTCCCTGAGTGGAGTGTCCGTTGCTGTG 3480
3481 GATCGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3540
3481 GATCGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3540
3541 GATTAATACGCAATGAGTAACTGAGTGAAGAGTTTCAATTCGCTTTCGCGGC 3600
3541 GATTAATACGCAATGAGTAACTGAGTGAAGAGTTTCAATTCGCTTTCGCGGC 3600
3601 CTTCAAGGTGACACTATTTGAGAGATTAAGTCAATATTAAGGCTCTTTTGGCGAATGT 3660
3601 CTTCAAGGTGACACTATTTGAGAGATTAAGTCAATATTAAGGCTCTTTTGGCGAATGT 3660
3661 ACAAAATCTTGCAGAAATCCCTTGTATCGGACAAATTAACAGTTAATATTTGTTAGC 3720
3661 ACAAAATCTTGCAGAAATCCCTTGTATCGGACAAATTAACAGTTAATATTTGTTAGC 3720
3721 TTTTGAACAAATTCATGCTGATATTTTGTGTTTCTTCCGCTTAAAGAAATTCAT 3780
3721 TTTTGAACAAATTCATGCTGATATTTTGTGTTTCTTCCGCTTAAAGAAATTCAT 3780
3781 GGAACATTAAGACCTCGCGCAACGCACTCTGCGCACATTTGGCGGAGAACATTTGT 3840
3781 GGAACATTAAGACCTCGCGCAACGCACTCTGCGCACATTTGGCGGAGAACATTTGT 3840
3841 CGCGCGCGCACCTGTGCAACGCGTTTACGCTGCTGCTCAAAAGACCAAGAGATGTGA 3900
3841 CGCGCGCGCACCTGTGCAACGCGTTTACGCTGCTGCTCAAAAGACCAAGAGATGTGA 3900
3901 TCGCCAAAGTCTGATGATGATCCAGATCTGAAAGGCACTTTTGAACCTGGGCGATGT 3960
3901 TCGCCAAAGTCTGATGATGATCCAGATCTGAAAGGCACTTTTGAACCTGGGCGATGT 3960
3961 CCAGATCAATCGTGGGCGAGGAGTGGATCACTTTTCAAAAGAACTGATGACGCAAC 4020
3961 CCAGATCAATCGTGGGCGAGGAGTGGATCACTTTTCAAAAGAACTGATGACGCAAC 4020
4021 CTTCAAGACATCGCTGTGCTCAAGACAGCTCAAGAGATGTTGTGCTTAAACGCGCA 4080
4021 CTTCAAGACATCGCTGTGCTCAAGACAGCTCAAGAGATGTTGTGCTTAAACGCGCA 4080
4081 CTGTTCAAGCGCTGCTGTGAAGATTTGGCGGACATTTTCTGCTCGCTGATCCATCTT 4140

XX 18-DEC-2000; 2000EP-00127688.
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX MPI; 2001-376931/40.
 XX
 PT Novel polynucleotides derived from *Corynebacterium* bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Disclosure; SEQ ID NO 7068; 246bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the *Corynebacterium* *Corynebacterium glutamicum*. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of *Corynebacterium* *Corynebacterium*, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC *Corynebacterium* *Corynebacterium*, and identifying a homologue of a gene derived from
 CC *Corynebacterium* *Corynebacterium*. *Corynebacterium* bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX

Sequence 349980 BP; 81250 A; 97718 C; 90621 G; 80391 T; 0 U; 0 Other;
 Query Match 94.8%; Score 5656.2; DB 5; Length 349980;
 Best Local Similarity 97.1%; Pred. No. 0;
 Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;

QY 1 AGTCGCTGACGCGCACCATTTGATGTGTGTGTCACGAGCTTGGCGAGGCTTTTACATCT 60
 DB 115585 AGTCGCTGACGCGCACCATTTGATGTGTGTGTCACGAGCTTGGCGAGGCTTTTACATCT 115526
 QY 61 ACCCTCCCGTGGCGGTGAGTGGGTCATTAACGGGTGGATCAGCCCGGTGAAAGTTGGC 120
 DB 115525 ACATCTCCGTCGCGGTGAGTGGGTCATTAACGGGTGGATCAGCCCGGTGAAAGTTGGC 115466
 QY 121 GAACCATGTGTCTCTGTGGGTGAGGGAAGAGTGGGGGAGAAAGTTTCAAGTG 180
 DB 115465 GAACCATGTGTCTCTGTGGGTGAGGGAAGAGTGGGGGAGAAAGTTTCAAGTG 115406
 QY 181 TCTGCAAGTTTAAAGTTATGATCATCAGCTTGAAGGCTGAGGTAAATTCAGTAGACTG 240
 DB 115405 TCTGCAAGTTTAAAGTTATGATCATCAGCTTGAAGGCTGAGGTAAATTCAGTAGACTG 115346
 QY 241 CAACAGAGGCTCAAGTCCGAAGATTAATTAACCTTGATCGTAGACATTAAGCATATA 300
 DB 115345 CAACAGAGGCTCAAGTCCGAAGATTAATTAACCTTGATCGTAGACATTAAGCATATA 115286
 QY 301 CGTCTATGCTTCTGGAAGGAACCAATTAACCTCAAGAAAGATGGAGAAAGTGTGCATT 360
 DB 115285 CGTCTATGCTTCTGGAAGGAACCAATTAACCTCAAGAAAGATGGAGAAAGTGTGCATT 115226
 QY 361 ATCAAGAAATGAGGTCAAGAGTTAAAAAATGAGGAGAAATTTTCCCTCCCTCG 420
 DB 115225 ATCAAGAAATGAGGTCAAGAGTTAAAAAATGAGGAGAAATTTTCCCTCCCTCG 115166
 QY 421 GGGTGAATGATGCTTTCTCAACTGAAAAGGCATCATCAGGAATCTTCTGAGAAC 480
 DB 115165 GGGTGAATGATGCTTTCTCAACTGAAAAGGCATCATCAGGAATCTTCTGAGAAC 115106
 QY 481 CACACCTAAAAAGGAGATTCACCCCGAAGCTCCCAAGATGTTCCCGGTTTATG 540

DB 115105 CAGCACTAAAAAGCAGGATTCACCCGAACTCCCAAGATGTGTCCAGTTTATG 115046
 QY 541 ATCTTCAATATACCGGTGAAAACGGTGGCGCTTCTTACCGGAACGAGACGAGCGCA 600
 DB 115045 ATCTTCAATATACCGGTGAAAACGGTGGCGCTTCTTACCGGAACGAGACGAGCGCA 114986
 QY 601 GGAACACCGCGGATATACCGCGAAACATGAGCAGACGAGTGTGTCACATATGTTT 660
 DB 114985 GGAATGCGCGGATATACCGCGAAACATGAGCAGACGAGTGTGTCACATATGTTT 114926
 QY 661 CGCGCGCGCTGACGCACTGGCAGCGAGGTGAAAACCTTATTCCTTGTGTAAGAG 720
 DB 114925 CGCGCGCGCTGACGCACTGGCAGCGAGGTGAAAACCTTATTCCTTGTGTAAGAG 114866
 QY 721 TCTGCTGTGCGCATTCATCTGAGGGCCCTTTATCAACGATCCGCTTGTGCTC 780
 DB 114865 GCTGCTGTGCGCATTCATCTGAGGGCCCTTTATCAACGATCCGCTTGTGCTC 114806
 QY 781 AAAACCCGGAATTCATTTTCCCGCAACCCAAAGATCTGCGCGGTGATCCATGCG 840
 DB 114805 AAAACCCGGAATTCATTTTCCCGCAACCCAAAGATCTGCGCGGTGATCCATGCG 114746
 QY 841 GAAAAGTTGATCAATTCATCAGTAGCGCGGAACTGACATCTTTCTGAGCTTC 900
 DB 114745 GAAAAGTTGATCAATTCATCAGTAGCGCGGAACTGACATCTTTCTGAGCTTC 114686
 QY 901 TCGATCTGCGCAGCGCACCATCATCTTCTTCTGCGGCACTGATGAGATTTTG 960
 DB 114685 TCGATCTGCGCAGCGCACCATCATCTTCTTCTGCGGCACTGATGAGATTTTG 114626
 QY 961 ATACCACTACGAGCGCAATGCTTGGCTAAAGAAAAATGACGCGTACGCGTACGC 1020
 DB 114625 ATACCACTACGAGCGCAATGCTTGGCTAAAGAAAAATGACGCGTACGCGTACGC 114566
 QY 1021 ATTTGTCATGCGATGCTTCCGCTGATCATATAGGCTCCCGAGCGTGGCGCTTTC 1080
 DB 114565 ATTTGTCATGCGATGCTTCCGCTGATCATATAGGCTCCCGAGCGTGGCGCTTTC 114506
 QY 1081 TTGCTGCGCAGTGGCGGGAACGATATGTTGATGATGCCCGACGCGCTGATTTGG 1140
 DB 114505 TTGCTGCGCAGTGGCGGGAACGATATGTTGATGATGCCCGACGCGCTGATTTGG 114446
 QY 1141 CCGATGGAACGGTCAATCTAGCTCGTCCCAACAGCCCTTTTCAACGAGCGCATGG 1200
 DB 114445 CCGATGGAACGGTCAATCTAGCTCGTCCCAACAGCCCTTTTCAACGAGCGCATGG 114386
 QY 1201 AAGCGCGCGAATGACAGCGGTGATCATTTTGGCGCTTTTGAACGTCAACGTCACCG 1260
 DB 114385 AAGCGCGCGAATGACAGCGGTGATCATTTTGGCGCTTTTGAACGTCAACGTCACCG 114326
 QY 1261 ATGAGTGGCGCGCTGCGGATGCGCGGCGCATGCGCGGGGACACAGACACATAGCGA 1320
 DB 114325 ATGAGTGGCGCGCTGCGGATGCGCGGCGCATGCGCGGGGATACAGACACATAGCGA 114266
 QY 1321 GTACGTTGTCACACGCGGCGGAGGTATGAGCTTATGACGCGACCTCCACACCT 1380
 DB 114265 GTACGTTGTCACACGCGGCGGAGGTATGAGCTTATGAGCTTATGACGCGACCTCCACACCT 114206
 QY 1381 CAACGCTGCGCTAAATTTCTCGGTCTTGGGATCAAGAAATGCTAAATCAACCTCG 1440
 DB 114205 CAACGCTGCGCTAAATTTCTCGGTCTTGGGATCAAGAAATGCTAAATCAACCTCG 114146
 QY 1441 CAAATTTGTGCTTTGACTCAAAAGCGCAGGTGCAAAAGTCCATTTAGTCAACAG 1500
 DB 114145 CAAATTTGTGCTTTGACTCAAAAGCGCAGGTGCAAAAGTCCATTTAGTCAACAG 114086
 QY 1501 TACTTAAAGTAAAGTAAATCTATCTGAT -TTTAAAGAGTCCCAACATGAAATCACT 1559
 DB 114085 TACTTAAAGTAAAGTAAATCTTCTGATTAATAAGAGTCCGACATGACATCATC 114026
 QY 1560 ATCTGCAAAAGACAGCAAGATCGCAAGAGTTCAGTCTAATGCAACCTTGGCG 1619
 DB 114025 ATCTGCAAAAGACAGCAAGATCGCAAGAGTTCAGTCTAATGCAACCTTGGCG 113966

QY 1620 AACAGGGTGAACCTTGGGCTTGGCAAGAGTCTCAACAAGTATCTTACCAAG 1679
Db 113965 ACTAAGGGGGGAACCTTGGGGCTTGGCACTGATCGTCACTTTTGGAGCACTTACCAAG 113906
QY 1680 CTCATTGCGATGATGAAGCTGGGGAAGTGTCACTTCAAGAGTGCAGAGCATTTCTGTG 1739
Db 113905 CTCATTGCGATGATGAAGCTGGGGAAGTGTCACTTCAAGAGTGCAGAGCATTTCTGTG 113846
QY 1740 GATGAATACGTGGGACTAAACCCGTGACGATGAAAAACGCTACTTTAAAAACATTCGCAA 1799
Db 113845 GATGAATACGTGGGACTTAAACGGCGACGATGAAAAACGCTACTTTAAAAACATTCGTA 113786
QY 1800 GAGTTCACGTACCAATCGACATCGTTGATGAAAGTGTCTAACGCCCAAGTGTGCAAC 1859
Db 113785 GAGTTCACGTACCAATCGACATCGTTGATGAAAGTGTCTAACGCCCAAGTGTGCAAC 113726
QY 1860 CCTGATCCATACGAAGCAGCTCAGAGTATGAGGCAAGATCGCTGCAAGATCCGTTGA 1919
Db 113725 CCTGATCCATACGAAGCAGCTCAGAGTATGAGGCAAGATCGCTGCAAGATCCGTTGAT 113666
QY 1920 GTTCAAAATCTTGGCATCGGCGGAAAAGG-CAATCGCTTTTCAATGAAACCATCATCTTCT 1978
Db 113665 GTTCAAAATCTTGGCATCGGCGGAAAAGG-CAATCGCTTTTCAATGAAACCATCATCTTCT 113606
QY 1979 CTGTGAGGACTGACAAAGGTCCAGGCGCTGCACCTTAAACGTGAGAGGACAAAGCTTGA 2038
Db 113605 CTGTGAGGACTGACAAAGGTCCAGGCGCTGCACCTTAAACGTGAGAGGACAAAGCTTGA 113546
QY 2039 TTCTTCAACACCATCGAAGAGGTCCCAACCCAGCGCTGCACCCAGAGTTTGGGCACTTGG 2098
Db 113545 TTCTTCAACACCATCGAAGAGGTCCCAACCCAGCGCTGCACCCAGAGTTTGGGCACTTGG 113486
QY 2099 TCCCGCGCGCAAAACATCGTGTGTGGCACTGTGTGAAGAAAAGCCGACCGCATCCGC 2158
Db 113485 TCCCGCGCGCAAAACATCGTGTGTGGCACTGTGTGAAGAAAAGCCGACCGCATCCGC 113426
QY 2159 GGAACGTGGAAGGCGCCAGTGACTGCTTGTGCCAGAGTTCCATCTGTGATGATGACAAAC 2218
Db 113425 GGAACGTGGAAGGCGCCAGTGACTGCTTGTGCCAGAGTTCCATCTGTGATGATGACAAAC 113366
QY 2219 AATGCAACCATCATCGTTGATGATGAAGAGAGCATCCAGCTGGAAGAGCTGTATCATAC 2278
Db 113365 AATGCAACCATCATCGTTGATGATGAAGAGAGCATCCAGCTGGAAGAGCTGTATCATAC 113306
QY 2279 CGTCTCATGAGACATTTAAAGCTGCGCTTGAAGAAAAGAAAAGTATCTGTGTGGGCTTA 2338
Db 113305 CGTCTCATGAGACATTTAAAGCTGCGCTTGAAGAAAAGAAAAGTATCTGTGTGGGCTTA 113246
QY 2339 TGCACACAGAACTTTCCAGTTTGGCGCCCTGCTACATGTGACTCTCCGCGAGGCGAGGC 2398
Db 113245 TGCACACAGAACTTTCCAGTTTGGCGCCCTGCTACATGTGACTCTCCGCGAGGCGAGGC 113186
QY 2399 TCAATGATCCAAAGGAAATGATCGTGAATGAGATACCTTCCACGCTTAACTACAGACG 2458
Db 113185 TCAATGATCCAAAGGAAATGATCGTGAATGAGATACCTTCCACGCTTAACTACAGACG 113126
QY 2459 ATCCAGGTTTCCCTTTCGACCAAGGCGACCGGCTGGGCTACACACACAGCCGTTGA 2518
Db 113125 ATCCAGGTTTCCCTTTCGACCAAGGCGACCGGCTGGGCTACACACAGCCGTTGA 113066
QY 2519 CCGGACCGGAGGATTTGAGTGAAGGACGACCTGCGCGACGCTCTTTAACCGGATGATCCT 2578
Db 113065 CCGGACCGGAGGATTTGAGTGAAGGACGACCTGCGCGACGCTCTTTAACCGGATGATCCT 113006
QY 2579 ATGACCTTGATGATGATGATTCGCGTGAAGCCGTAATTTAATGACGCAACTTAAACTTT 2638
Db 113005 ATGACCTTGATGATGATGATTCGCGTGAAGCCGTAATTTAATGACGCAACTTAAACTTT 112946
QY 2639 TCTTACACCGGCACTTAAATAATTTGACGAAAGCGCGGCGCACCCAAAACCTTGTGCAAG 2698
Db 112945 TCTTACACCGGCACTTAAATAATTTGACGCGGCGCGCGCACCCAAAACCTTGTGCAAG 112886

QY 2699 TCGAGGACCCCAACTGGGCTGATGGGGCGGCACTTCAATCGCGCTTAAATAATCGGCTTA 2758
Db 112885 TCGAGGACCCCAACTGGGCTGATGGGGCGGCACTTCAATCGCGCTTAAATAATCGGCTTA 112826
QY 2759 TCGACGAGCCCGCAGCGGTTTACACCCCACTTACCGGATGCCATGATACAGCCCTGATG 2818
Db 112825 TCGACGAGCCCGCAGCGGTTTACACCCCACTTACCGGATGCCATGATACAGCCCTGATG 112766
QY 2819 GTGATGTTGGAACATGCTTCTTGGGGCCCAACCGGAAAACCTCAACCGGTGACGCGTTC 2878
Db 112765 GTGATGTTGGAACATGCTTCTTGGGGCCCAACCGGAAAACCTCAACCGGTGACGCGTTC 112706
QY 2879 TATACCGCTGCACAGATCTTGAATACTGGGAATCTCGGGAATCACTTTGACCTCA 2938
Db 112705 TATACCGCTGCACAGATCTTGAATACTGGGAATCTCGGGAATCACTTTGACCTCA 112646
QY 2939 GTGATGCAAACTGTGTCTGCTCTGATCTCGTTCCGATGAGCTACATGTGGGAATGCC 2998
Db 112645 GTGATGCAAACTGTGTCTGCTCTGATCTCGTTCCGATGAGCTACATGTGGGAATGCC 112586
QY 2999 CCAACCTTTTAAACGCTTTCGATGAAGAAACCTGGGAAATCTGACGCTGATTTTCT 3058
Db 112585 CCAACCTTTTAAACGCTTTCGATGAAGAAACCTGGGAAATCTGACGCTGATTTTCT 112526
QY 3059 GTCCACAAAGATGGAACCGAATCCACGATGAGGTTACTCATACGCAAGCTCTGACAGT 3118
Db 112525 GTCCACAAAGATGGAACCGAATCCACGATGAGGTTACTCATACGCAAGCTCTGACAGT 112466
QY 3119 GCGATATGTGTGACAAAGCTTGAAGAAAGCACTTCCGCGTCTTGGAGATTCAGCG 3178
Db 112465 GCGATATGTGTGACAAAGCTTGAAGAAAGCACTTCCGCGTCTTGGAGATTCAGCG 112406
QY 3179 AGCTGGAATTTGGGCTCATGAAATTTTACGCAACCGAGGTTGAGTAAAGGTTCTGATGCT 3238
Db 112405 AGCTGGAATTTGGGCTCATGAAATTTTACGCAACCGAGGTTGAGTAAAGGTTCTGATGCT 112346
QY 3239 GGCCTGAGGCTGATGAGGCGCTGCCCGCAGAGTATCAACCAAGTGTCAAGAGAG 3298
Db 112345 GGCCTGAGGCTGATGAGGCGCTGCCCGCAGAGTATCAACCAAGTGTCAAGAGAG 112286
QY 3299 GATGGTGCACCTGCTGACTGTGCTGCCCGCAAGCTTCAATTTGCGAACCGACGATCTAC 3358
Db 112285 GATGGTGCACCTGCTGACTGTGCTGCCCGCAAGCTTCAATTTGCGAACCGACGATCTAC 112226
QY 3359 AAGAGCTCTCTTCCCAAGGAGGAGTGGGGGTAAACGATCTGTATTTAGTCTTGAC 3418
Db 112225 AAGAGCTCTCTTCCCAAGGAGGAGTGGGGGTAAACGATCTGTATTTAGTCTTGAC 112166
QY 3419 CTGTCCGATGACATCCGAGGCAATTTTCCCTCGAGTGGAGTGTGTCCGTTGTCTG 3478
Db 112165 CTGTCCGATGACATCCGAGGCAATTTTCCCTCGAGTGGAGTGTGTCCGTTGTCTG 112106
QY 3479 TGGATCGTATGATGATGATGCTGCGGTAGTGAAGTAAACCTGGCGAATTAATGATCGCG 3538
Db 112105 TGGATCGTATGATGATGATGCTGCGGTAGTGAAGTAAACCTGGCGAATTAATGATCGCG 112046
QY 3539 ACGATTAATACGACATGAGTAACTGACAGTGTGACAGGTTTCAATGCTTTTCCG 3598
Db 112045 ACGATTAATACGACATGAGTAACTGACAGTGTGACAGGTTTCAATGCTTTTCCG 111987
QY 3599 GCGTTCAAAGGTGACACTATTTGAGAGATAGTCAATTAAGAGGCTCTTTTGTGGCAATT 3658
Db 111986 GCGTTCAAAGGTGACACTATTTGAGAGATAGTCAATTAAGAGGCTCTTTTGTGGCAATT 111927
QY 3659 GTACAAATATCTTCCCAAAATCTCTTGAATCGGACAAATTAACAGTTTATGTGTTTA 3718
Db 111926 GTACAAATATCTTCCCAAAATCTCTTGAATCGGACAAATTAACAGTTTATGTGTTTA 111867
QY 3719 GCTTTTGAACAATTCATGATCTGATATTTTGTGTTTCTTCCCGTTAAGGAAATTC 3778
Db 111866 GCTTTTGAACAATTCATGATCTGATATTTTGTGTTTCTTCCCGTTAAGGAAATTC 111807
QY 3779 ATGAGACATTAAGGACTTCGCGCAACGATCCTGCGGACATTTGGCGGCAAGCAACTT 3838

Db	111806	ATGACATTAAGAACTTGGCGAAGCCATCTTGGCGAATTGGCCGCGAAGCAACATT	111747
Oy	3839	GTGCGCCCGCCGACACTGTGCAACGCGTTTACGCTTGTGCTCCAAAGACACCAAGATGTG	3898
Db	111746	GTGCGCGCGGACACCTGTGCAACGCGTTTACGCTTGTGCTCCAAAGACACCAAGATGTG	111687
Oy	3899	GATCGCCAAAGTCTGGATGATGATCCAGATCTGAAGGCACTTTGAACCTGGCGCATG	3958
Db	111686	GATCGCCAAAGTCTGGATGATGATCCAGATCTGAAGGCACTTTGAACCTGGCGCATG	111627
Oy	3959	TTCCAGATCATGTGCGGCGCAGGCGCATGTGAGATCATGTTTCCAAAGACCTGCATGACGA	4018
Db	111626	TTCCAGATCATGTGCGGCGCAGGCGCATGTGAGATCATGTTTCCAAAGACCTGCATGACGA	111567
Oy	4019	ACCTCCAAAGACATGCTGTGTGTCCACAGACAGCTCAAGATGTTGTGCTAACACGCG	4078
Db	111566	ACCTCCAAAGACATGCTGTGTGTCCACAGAGAGCTCAAGATGTTGTGCTAACACGCG	111507
Oy	4079	AACTGGTTACGCCGTGCTGTGAAGGTATTTGGCGGACATTTTCGTCGCTGATTCGAATC	4138
Db	111506	AACTGGTTACGCCGTGCTGTGAAGGTATTTGGCGGACATTTTCGTCGCTGATTCGAATC	111447
Oy	4139	TTGGTTGGTGGGGGTCGTGCTCATGCTATCACAAGTGTGTGTGGCCAGATCTGTTC	4198
Db	111446	TTGGTTGGTGGGGGTCGTGCTCATGCTATCACAAGTGTGTGTGGCCAGATCTGTTC	111387
Oy	4149	GGTCCGCAATCACTGGGTGAGATGTTCCCTCCAGATCAGCGGTGTTCGAGATGATCAAC	4258
Db	111386	GGTCCGCAATCACTGGGTGAGATGTTCCCTCCAGATCAGCGGTGTTCGAGATGATCAAC	111327
Oy	4259	CTGATGGCATCTGCGCGGTTCGCGTTCCTTCCAGATGTGTGTTCACCGCAACGA	4318
Db	111326	CTCATGGCATCTGCGCGGTTCGCGTTCCTTCCAGATGTGTGTTCACCGCAACGA	111267
Oy	4319	CGTTTCGGTGGCAATGAGTTCCTGGGCGCGCGCATTTGATGGCATGTGTTCACACC	4378
Db	111266	CGTTTCGGGCGGCAATGAGTTCCTGGGCGCGCGCATTTGATGGCATGTGTTCACACC	111207
Oy	4379	CTGGTTAACGCGTTACGACGTTGCGCGCACATGACCGCGGCGCAATGTGCTGC	4438
Db	111206	TTGGTGAAACGCGTTACGACGTTGCGCGCACATGCTGCGGCGCAATGTGCTGC	111147
Oy	4439	CTGTTTGGTTGGATGTGTCTCAAGCTGTGTTACAGGCGACCGTGTCTGTGCTGATG	4498
Db	111146	CTGTTTGGTTGGATGTGTCTCAAGCTGTGTTACAGGCGACCGTGTCTGTGCTGATG	111087
Oy	4499	GTCCTCTTGATCTGTGCAACGATCCAGAAATTCCTGCAACAGGACTCATGGCACTGCA	4558
Db	111086	GTCCTCTTGATCTGTGCAACGATCCAGAAATTCCTGCAACAGGACTCATGGCACTGCA	111027
Oy	4559	GACCTTCGATCAACCCAGGTGTGACTGTGCTGCTACACGCGCTTCCTTACGTTCAATGCT	4618
Db	111026	GACCTTCGATCACTTCAGGTGTGACTGTGCTGCTACACGCGCTTCCTTACGTTCAATGCT	110967
Oy	4619	ATTGTGTCCAGCAATGCGCTGGGTGGGTGACTTGTCTGGCACACGCTGTGACGGGACTCTAT	4678
Db	110966	ATTGTGTCCAGCAATGCGCTGGGTGGGTGACTTGTCTGGCACACGCTGTGACGGGACTCTAT	110907
Oy	4679	GATTTTCGGTGGTCCAGTCCGCGGTCTGCTTTTCGCTGTGCTACTACCAATCGTTATC	4738
Db	110906	GATTTTCGGTGGTCCAGTCCGCGGTCTGCTTTTCGCTGTGCTACTACCAATCGTTATC	110847
Oy	4739	ACTGGTTCGACACAGTCCCTCCCGCAATTGAGCTGGAACCTGTTCACACAGGTGGATCC	4798
Db	110846	ACTGGTTCGACACAGTCCCTCCCGCAATTGAGCTGGAACCTGTTCACACAGGTGGATCC	110787
Oy	4799	TTTCATCTTCGCAACGCGCATCTCATGTGCGCAATATTCGCGAGGTGTGACATGTTTGGCAGTG	4858
Db	110786	TTTCATCTTCGCAACGCGCATCTCATGTGCGCAATATTCGCGAGGTGTGACATGTTTGGCAGTG	110727
Oy	4859	TTCTTCTTCAGCAAGATGAAAAAGCTCAAGGCGCTTTCAGGTGCTTCAAGTGTCTTCGCT	4918

Db	110726	TTCTTCCTGCGAAGAGTGAAGAAAGCTCAAGGGCCTTGACAGGTGCTTCAGGTGCTCCGCT	110661
Qy	4919	GTTCTTGGTATTACAGAGCTGCGACTCTTCCGTGTGAACCTTCCGCTGCGTGGCGGTTT	4978
Db	110666	GTTCTTGGTATTACAGAGCTGCGACTCTTCCGTGTGAACCTTCCGCTGCGTGGCGGTTT	110607
Qy	4979	TACATTGGTATCCGTAACCGACGCTATCCGTGGCGCTTTGATTGCACTCTTTGATATCAAG	5038
Db	110606	TTCAATCCGATATCGTATCCGACGACCTATCCGTGGCGCTTTGATTGCACTCTTTAATATCAAG	110547
Qy	5039	GCAGTTGGGTTGGGGGCTGCGAGGTTCTTGGGTGTTGTTCTTATGATGCTCCAGATATG	5098
Db	110546	GCAGTTGGGTTGGGGGCTGCGAGGTTCTTGGGTGTTGTTCTTATGATGCTCCAGATATG	110487
Qy	5099	GTCAATGTTCTTGGTTTGGCGCGGTAGTACCTTTGTATCCGATTCGGCGCAGCGATTTGCT	5158
Db	110486	GTCAATGTTCTTGGTTTGGCGCGGTAGTACCTTTGTATCCGATTCGGCGCAGCGATTTGCT	110427
Qy	5159	TATGGCCTTTACTTGGTTGCTGCCGCAACCGCAGACATTGATCCAGATGCAACCGTGCTCCA	5218
Db	110426	TATGGCCTTTACTTGGTTGCTGCCGCAACCGCAGACATTGATCCAGATGCAACCGTGCTCCA	110367
Qy	5219	GTGCGCTGAGGAAGCAACCAAGCCGAGAGAGAACCCCGAGATTTTTCAAACGATTC	5278
Db	110366	GTGCGCTGAGGAAGCAACCAAGCCGAGAGAGAACCCCGAGATTTTTCAAACGATTC	110307
Qy	5279	ACCATCATTCAGAGCACTTTGACCGGTGAGCTATCCGACTAGCAGCTGACGATGCC	5338
Db	110306	ACCATCATTCAGAGCACTTTGACCGGTGAGCTATTCGACTAGCAGCTGACGATGCC	110247
Qy	5339	ATGTTTGCACGCGGAAGCTTGGCTCAGGTGTTGCGATGTCGCCAACCAAGGGGACGCTG	5398
Db	110246	ATGTTTGCACGCGGAAGCTTGGCTCAGGTGTTGCGATGTCGCCAACCAAGGGGACGTTA	110187
Qy	5399	GTTTTCACAGTGAAGCGGAAAGATCGTGTGGGCTTCCCATCTGTGTCAGGCTTTCCGAGTC	5458
Db	110186	GTTTTCACAGTGAAGCGGAAAGATCGTGTGGGCTTCCCATCTGTGTCAGGCTTTCCGAGTT	110127
Qy	5459	CGCATAGAGCTGAGAGATGTTTCCATGTGATCTTGTGATGACACATTTGTTTCGACACC	5518
Db	110126	CGCATAGAGCTGAGAGATGTTTCCATGTGATCTTGTGATGACACATTTGTTTCGACACA	110067
Qy	5519	GTAACCTTCACCGGACGCACTTTAACCCGCTGAAGAAACGAGGCGATGACGTAAGCA	5578
Db	110066	GTAACCTTCACCGGACGCACTTTAACCCGCTGAAGAAACGAGGCGATGACGTAAGCA	110007
Qy	5579	GGGAGCGTGTGTGTGATTTGATTTGATGCACTTAAGGCTGCGAGGTTATAGGTAAAC	5638
Db	110006	GGGAGCGTGTGTGTGATTTGATTTGATGCACTTAAGGCTGCGAGGTTATAGGTAAAC	109947
Qy	5639	ACGCCGATTTGTTGTTTGCATTAACAAGAAACCCGACCTGTAAACATTAACGTTTGGGC	5698
Db	109946	ACGCCGATTTGTTGTTTGCATTAACAAGAAACCCGACCTGTAAACATTAACGTTTGGGC	109887
Qy	5699	GAATTTGAAGCGGAGCAACCTGTCTCAACGTCCGAAAGAAAGAACGGGTGCGACACACA	5758
Db	109886	GAATTTGAAGCGGAGCAACCTGTCTCAACGTCCGAAAGAAAGAACGGGTGCGACACACA	109827
Qy	5759	CCATTAAGTTGAAACCTTGAAGTGTGACACAGAGTTTGAATTAAGGGAAGCTGACTTACCC	5818
Db	109826	CCATTAAGTTGAAACCTTGAAGTGTGACACAGAGTTTGAATTAAGGGAAGCTGACTTACCC	109767
Qy	5819	ATCTTTGACACCGGATACCGGTAACGCTTGAGATTTTAAACCTGTTCAACCAAGTCAAGCC	5878
Db	109766	ATCTTTGACACCGGATACCGGTAACGCTTGAGATTTTAAACCTGTTCAACCAAGTCAAGCC	109707
Qy	5879	TCCGTGTACCTGTGTGTGTCACACCCCGCAATCTTCAACCCCATTTGACATGTTGCTTCA	5938
Db	109706	TCCGTGTACCTGTGTGTGTCACACCCCGCAATCTTCAACCCCATTTGACATGTTGCTTCA	109647
Qy	5939	GCAGTAGCGTTGATATTTTGGCGCGCTG	5967
Db	109646	GCAGTAGCGTTGATATTTTGGCGCGCTG	109618

Db 961 ATACACATACAGCGCAATTCCTTGCTAAAGAAAATGAGCGTCAACGCCCTACCG 1020
Qy 1021 ATTTGTTCAATGCGATGCTCTCCGCTGATATAGGGCTCCGGACGCTGGCGCTTTGC 1080
Db 1022 ATTTGTTCAATGCGATGCTCTCCGCTGATATAGGGCTCCGGACGCTGGCGCTTTGC 1080
Qy 1081 TTGCTCGGACGCTGGCGGAGCGCATGTGATGATTCGCGGACGGGCTGATTTGG 1140
Db 1081 TTGCTCGGACGCTGGCGGAGCGCATGTGATGATTCGCGGACGGGCTGATTTGG 1140
Qy 1141 CGGATGGAACGGTGAATCTAGTCTGTCCAAACAACCCCTTTTTCATCAGGAGCCATGG 1200
Db 1141 CGGATGGAACGGTGAATCTAGTCTGTCCAAACAACCCCTTTTTCATCAGGAGCCATGG 1200
Qy 1201 AAGCCGCCGGAATGCCAGACGGTGAATCAATTTGGGCGCTTTTGAACGTCAACCTGACCG 1260
Db 1201 AAGCCGCCGGAATGCCAGACGGTGAATCAATTTGGGCGCTTTTGAACGTCAACCTGACCG 1260
Qy 1261 ATGAGATCGCCGCTGCGCGGATGGGGGCGCATTCGCCGGGGGACACAGACATAGCGA 1320
Db 1261 ATGAGATCGCCGCTGCGCGGATGGGGGCGCATTCGCCGGGGGACACAGACATAGCGA 1320
Qy 1321 GTCAAGTTCGTGACACAGTGGCGAGGGGTATGACGCTTATCGACGGACCTTCACACT 1380
Db 1321 GTCAAGTTCGTGACACAGTGGCGAGGGGTATGACGCTTATCGACGGACCTTCACACT 1380
Qy 1381 CAACCGTGGCGCTTAAATTTCTGCTTTGGCGATACGAAATCGCTAAATCCAAACCTTG 1440
Db 1381 CAACCGTGGCGCTTAAATTTCTGCTTTGGCGATACGAAATCGCTAAATCCAAACCTTG 1440
Qy 1441 CAAATTTGAGCTTTGACTGCTCAACGCGGACGGTGAAGGATCACTTAAGGATCAAG 1500
Db 1441 CAAATTTGAGCTTTGACTGCTCAACGCGGACGGTGAAGGATCACTTAAGGATCAAG 1500
Qy 1501 TACTTTAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
Db 1501 TACTTTAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
Qy 1561 TCTGCAAAAGACGAGAAAGTGGCAAAAGCACTTCAAGATCGTCAATCGACCTTGGCA 1620
Db 1561 TCTGCAAAAGACGAGAAAGTGGCAAAAGCACTTCAAGATCGTCAATCGACCTTGGCA 1620
Qy 1621 ACAAGGGTGAACCTTGAGGGCTTGGCAACGAGATCTCACACGATACCTTACCAAGAC 1680
Db 1621 ACAAGGGTGAACCTTGAGGGCTTGGCAACGAGATCTCACACGATACCTTACCAAGAC 1680
Qy 1681 TCAATTCGATATGAGAGTGGGGAAGTGTCAATCAAGAACTGCAAGCATTCCTTGG 1740
Db 1681 TCAATTCGATATGAGAGTGGGGAAGTGTCAATCAAGAACTGCAAGCATTCCTTGG 1740
Qy 1741 ATGAATAGTGGGATCAACCCGTGACGATGAAAACAGCTACTTTAAACCATTTGCAAG 1800
Db 1741 ATGAATAGTGGGATCAACCCGTGACGATGAAAACAGCTACTTTAAACCATTTGCAAG 1800
Qy 1801 AGTTACATGACCAATCGATCATGTGATGAAAGAGTCTACAGCCGAGATGGGCAAC 1860
Db 1801 AGTTACATGACCAATCGATCATGTGATGAAAGAGTCTACAGCCGAGATGGGCAAC 1860
Qy 1861 CTGATTCATACGAGAGCTGACAGATATAGGCAAAAGATCGCTGACAGATCGTGAAG 1920
Db 1861 CTGATTCATACGAGAGCTGACAGATATAGGCAAAAGATCGCTGACAGATCGTGAAG 1920
Qy 1921 TTCAAAATCTTGGCATCGCGGAAACGGCAATCGCTTTCAATTTGAACATCATCTTCT 1980
Db 1921 TTCAAAATCTTGGCATCGCGGAAACGGCAATCGCTTTCAATTTGAACATCATCTTCT 1980
Qy 1981 GTCAAGACTGACAAAGGCTCAAGGCGCTGACCTTAAACTGTGAGAGCAACGCTGAT 2040
Db 1981 GTCAAGACTGACAAAGGCTCAAGGCGCTGACCTTAAACTGTGAGAGCAACGCTGAT 2040
Qy 2041 GTTCAACACATGAAAGGTCCCAACCAACGCGCTCAACGAGGTTTGGGCACTTTGTC 2100
Db 2041 GTTCAACACATGAAAGGTCCCAACCAACGCGCTCAACGAGGTTTGGGCACTTTGTC 2100

Qy 2101 CCGCGCGCAAAACATCGTGTGGTGCACTGTGAAGAAAAGCCGACCTCCGCGG 2160
Db 2101 CCGCGCGCAAAACATCGTGTGGTGCACTGTGAAGAAAAGCCGACCTCCGCGG 2160
Qy 2161 AACTGGAAGGCGCGAGTACGTCTTCTGGCCAGGTTTCATCTGTAGATCACAAAT 2220
Db 2161 AACTGGAAGGCGCGAGTACGTCTTCTGGCCAGGTTTCATCTGTAGATCACAAAT 2220
Qy 2221 GCCACCATCATGTGTGATGAAAGCAGATATCAAGCTGAAAACGCTGATCACTACCG 2280
Db 2221 GCCACCATCATGTGTGATGAAAGCAGATATCAAGCTGAAAACGCTGATCACTACCG 2280
Qy 2281 TCTCATGAGCAATTAAGCTGCGCTGAAAACAAAAGGAAAGTCTGTGGGCTATG 2340
Db 2281 TCTCATGAGCAATTAAGCTGCGCTGAAAACAAAAGGAAAGTCTGTGGGCTATG 2340
Qy 2341 CACACGAACTTCCAGTTTGGCGCTGCTGCAATGTAACCTCCGAGGGCAGGCTC 2400
Db 2341 CACACGAACTTCCAGTTTGGCGCTGCTGCAATGTAACCTCCGAGGGCAGGCTC 2400
Qy 2401 AATGATCCCAACGGAATGTACGTGATGAGATACCTTCACGCTACTACAGACAT 2460
Db 2401 AATGATCCCAACGGAATGTACGTGATGAGATACCTTCACGCTACTACAGACAT 2460
Qy 2461 CCAAGTTTCCCTTGGCACAAAGCGCACCGGCTGGGCTCACACCAACGCGCTTGAC 2520
Db 2461 CCAAGTTTCCCTTGGCACAAAGCGCACCGGCTGGGCTCACACCAACGCGCTTGAC 2520
Qy 2521 GGAACGAGCGATGAGTGAAGTGAAGCAGCACTGCGCGACGCTCTTACCCGAGTGCAT 2580
Db 2521 GGAACGAGCGATGAGTGAAGTGAAGCAGCACTGCGCGACGCTCTTACCCGAGTGCAT 2580
Qy 2581 GACCTGATGATGCTATTCGCTGAGAGCGGATTTACTGACGGAACATTAACCTTTTC 2640
Db 2581 GACCTGATGATGCTATTCGCTGAGAGCGGATTTACTGACGGAACATTAACCTTTTC 2640
Qy 2641 TACACCGGCAACTTAAATTTGACGAAAGCGCGCGCCACCAAAACCTTGTGCAAGTC 2700
Db 2641 TACACCGGCAACTTAAATTTGACGAAAGCGCGCGCCACCAAAACCTTGTGCAAGTC 2700
Qy 2701 GAGGACCCCAACTGGGCTGATGGGCGGCAATTCATGCGCGTGTGCTTAAATCCGCTATC 2760
Db 2701 GAGGACCCCAACTGGGCTGATGGGCGGCAATTCATGCGCGTGTGCTTAAATCCGCTATC 2760
Qy 2761 GACGAGCCGCGAGCGGTTTCAACCCCATTTACCGGATCCCATGATGAGCCCTGATGT 2820
Db 2761 GACGAGCCGCGAGCGGTTTCAACCCCATTTACCGGATCCCATGATGAGCCCTGATGT 2820
Qy 2821 GATGTTGGAACATGCTTCTGGGCGCCACCGGAAAACCTTACCGGCTGACCGCTTCTA 2880
Db 2821 GATGTTGGAACATGCTTCTGGGCGCCACCGGAAAACCTTACCGGCTGACCGCTTCTA 2880
Qy 2881 TACCGCTGACGATCTTGA AAACTGGGAAATCTCCGCGGAAATCACTTTGACCTCACT 2940
Db 2881 TACCGCTGACGATCTTGA AAACTGGGAAATCTCCGCGGAAATCACTTTGACCTCACT 2940
Qy 2941 GATGCAAACTGTGTTCTGCTCTGATCTGTTCCGATGAGCTACATGATGGAATGCCCC 3000
Db 2941 GATGCAAACTGTGTTCTGCTCTGATCTGTTCCGATGAGCTACATGATGGAATGCCCC 3000
Qy 3001 AACCTTTTTCGCTTCCGATGAGAAATCGGCGAAATCTGACGTGTGATTTTCTGT 3060
Db 3001 AACCTTTTTCGCTTCCGATGAGAAATCGGCGAAATCTGACGTGTGATTTTCTGT 3060
Qy 3061 CCAAGAGATTGACCGAATCCAGATGAGTTACTCTACGCAAGCTTGAACCAAGTGC 3120
Db 3061 CCAAGAGATTGACCGAATCCAGATGAGTTACTCTACGCAAGCTTGAACCAAGTGC 3120
Qy 3121 GGAATATGCTGTCGACAACTTGAAGAAAGCACTTCCGCGCTTTCGAGATTCAGCAG 3180
Db 3121 GGAATATGCTGTCGACAACTTGAAGAAAGCACTTCCGCGCTTTCGAGATTCAGCAG 3180

QY 3181 CTGATTTTCGGCATGAAATTCTACGACCGAGTTGCAATACGTTTCTGATCCCTGG 3240
 DB 3181 CTGATTTTCGGCATGAAATTCTACGACCGAGTTGCAATACGTTTCTGATCCCTGG 3240
 QY 3241 CTCCTGGGCTGATGAGGGGTGCGCCCGGAGATGATCACCAGTTGCAAGAGAGA 3300
 DB 3241 CTCCTGGGCTGATGAGGGGTGCGCCCGGAGATGATCACCAGTTGCAAGAGAGA 3300
 QY 3301 TGGGTGCACTGCTGACTGTGCGCCGCAAGCTTCAATTGCGCAACACGCGATCTACCA 3360
 DB 3301 TGGGTGCACTGCTGACTGTGCGCCGCAAGCTTCAATTGCGCAACACGCGATCTACCA 3360
 QY 3361 GAGCTCTTCTCCAGAGGGGGAGTGGGGGTAATCAAGTCTGTATTAAGTTCTAACT 3420
 DB 3361 GAGCTCTTCTCCAGAGGGGGAGTGGGGGTAATCAAGTCTGTATTAAGTTCTAACT 3420
 QY 3421 GTCCGAGTGAATCCGAGGCAATATTTCCCTGAGTGGAGTGTCCGTTTGTCTGTG 3480
 DB 3421 GTCCGAGTGAATCCGAGGCAATATTTCCCTGAGTGGAGTGTCCGTTTGTCTGTG 3480
 QY 3481 GATCGTATGATGATCGTGCCTAGCTAGTAAACCTGCGCAATTAGTATCGCGAC 3540
 DB 3481 GATCGTATGATGATCGTGCCTAGCTAGTAAACCTGCGCAATTAGTATCGCGAC 3540
 QY 3541 GATTAATACGCCAATTGAGTAACTGAGGTATGACAGGTTTCAATTCCTTTCCGGGC 3600
 DB 3541 GATTAATACGCCAATTGAGTAACTGAGGTATGACAGGTTTCAATTCCTTTCCGGGC 3600
 QY 3601 CTTCAGAAAGTGAACATTAAGAGATTAAGTCAATTAAGAGTCTTTTGTGCGCAATTGT 3660
 DB 3601 CTTCAGAAAGTGAACATTAAGAGATTAAGTCAATTAAGAGTCTTTTGTGCGCAATTGT 3660
 QY 3661 ACAAAATAGCTTCGCAAAATCCCTTGAATC 3687
 DB 3661 ACAAAATAGCTTCGCAAAATCCCTTGAATC 3687
 RESULT 4
 ID AAH67869 standard; DNA; 1983 BP.
 AC AAH67869;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 2904.
 XX
 KM Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 OS
 PN BP1108790-A2.
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000BP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI: 2001-376931/40.
 DR P-PSDB; AAG92650.
 XX
 PT Novel polynucleotides derived from Corynebacterium, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.

XX
 PS Claim 8; SEQ ID NO 2904; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium, and identifying a homologue of a gene derived from
 CC corynebacterium. Corynebacterium are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 1983 BP; 410 A; 508 C; 558 G; 507 T; 0 U; 0 Other;
 Query Match 31.4%; Score 1874.2; DB 5; Length 1983;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1915; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 3779 ATGAGCATAAAGACCTGCGCAACGATCTGCGCAATTTGGCGGCAAGACATTT 3838
 DB 1 ATGAGCATAAAGACCTGCGCAACGATCTGCGCAATTTGGCGGCAAGACATTT 60
 QY 3839 GTCCCGCCGCACTGTGCAACGCCGTTTACGCTTCGTGTCGCAAGACCAAGAGATGTG 3898
 DB 61 GTCCCGCCGCACTGTGCAACGCCGTTTACGCTTCGTGTCGCAAGACCAAGAGATGTG 120
 QY 3899 GATGCGCAAAAGTCTGAGTATGATGATCCAGATCTGAAAGGCACTTTGAAACTGGCGCATG 3958
 DB 121 GATGCGCAAAAGTCTGAGTATGATGATCCAGATCTGAAAGGCACTTTGAAACTGGCATG 180
 QY 3959 TTCCAGATCATGCTGCGGCCGCGCGATGATGATGATTTTCAAGAACTGATGAGCGCA 4018
 DB 181 TTCCAGATCATGCTGCGGCCGCGCGATGATGATGATTTTCAAGAACTGATGAGCGCA 240
 QY 4019 AACTCCAAAGACATGCGCTGTGTCACAGAGAGCTCAAGAGTGTGGGCTTAACAACGCC 4078
 DB 241 AACTCCAAAGACATGCGCTGTGTCACAGAGAGCTCAAGAGTGTGGGCTTAACAACGCC 300
 QY 4079 AACTGTTTCAGCCGCTGTGTAAGGTAATGGCGGCAATTTTGTCCGCTGATTCGAATC 4138
 DB 301 AACTGTTTCAGCCGCTGTGTAAGGTAATGGCGGCAATTTTGTCCGCTGATTCGAATC 360
 QY 4139 TTGGTTGGTGGCGGTCTGCTCATGCTTCAACAATGTTGGTGGCGGAGATCTGTTTC 4198
 DB 361 TTGGTTGGTGGCGGTCTGCTCATGCTTCAACAATGTTGGTGGCGGAGATCTGTTTC 420
 QY 4199 GGTCCGCAATCACTGTGGAAGATGTTCCCTCAGATCAGCGGTGTTGCTGATGATCAAC 4258
 DB 421 GGTCCGCAATCACTGTGGAAGATGTTCCCTCAGATCAGCGGTGTTGCTGATGATCAAC 480
 QY 4259 CTGATGATCATGCGCCGTTGCGCTTCTTGGCAAGTGTGTTGTTTCAACCGCAACAG 4318
 DB 481 CTGATGATCATGCGCCGTTGCGCTTCTTGGCAAGTGTGTTGTTTCAACCGCAACAG 540
 QY 4319 CGTTTCGTTGCGCAATGATGTTCTTGGGCGCGGCAATGTTGTAAGGAGTGTGTTCCAGC 4378
 DB 541 CGTTTCGTTGCGCAATGATGTTCTTGGGCGCGGCAATGTTGTAAGGAGTGTGTTCCAGC 600
 QY 4379 CTGTTTAAAGGCTACGAGTGGCGGCGCAATGACCGGCGGCAATGCAATGTGTCTC 4438
 DB 601 TTGGTGAACGCTACGAGTGGCGGCGCAATGACCGGCGGCAATGCAATGTGTCTC 660
 QY 4439 CTGTTTGGTTGATGTTGCTCAAGCTGTTTCAAGGCGACCGTGTCTTCTGTGTGTG 4498
 DB 661 CTGTTTGGTTTGAAGTGTGCGCAAGCCGTTTCAAGGCGACCGTGTCTTCTGTGTGTG 720
 QY 4499 GTCTCTTGAATCTGAGCAAGATGAGAAAGTTCCTGCAAGGCACTGATGAGGCACTGCA 4558
 DB 721 GTCTCTTGAATCTGAGCAAGATGAGAAAGTTCCTGCAAGGCACTGATGAGGCACTGCA 780

```

QY 4559 GACTTCCTGATGACCCAGAGTTGAAGTCTGCTGCTCAAGGGCTTCTTACGTTCAATGCT 4618
DB 781 GACTTCCTGATCACTCAGAGTGTGACGTTGCTGCTCAAGGATCTTTCATTCATCGCC 840
QY 4619 ATTGGTCAGCAATGCGTGGTGGTGAAGTCTGCTGGCAACAGTCTGAGGAACTCTAT 4678
DB 841 ATTGGCCAGCAATGGCTGGTGGTGGTGAAGTCTGCTGGCAACAGTCTGAGGAACTCTAT 900
QY 4679 GATTTCGGTGTGCTCAAGTGGCGGTCTGCTTTTCGGTCTGCTCTACTACCAATCGTTATC 4738
DB 901 GATTTCGGTGTGCTCAAGTGGCGGTCTGCTTTTCGGTCTGCTCTACTACCAATCGTTATC 960
QY 4739 ACTGCTGTCAGCAAGTCTTCCCGCAATTGAGCTGGAGCTGTTCACCAAGGTCGATCC 4798
DB 961 ACTGCTGTCAGCAAGTCTTCCCGCAATTGAGCTGGAGCTGTTCACCAAGGTCGATCC 1020
QY 4799 TTTCATCTTCGCAACCGCATCTCATGAGCAATATCGCGAGGTCAGCATGTTTGGCAGTG 4858
DB 1021 TTTCATCTTCGCAACCGCATCTCATGAGCAATATCGCGAGGTCAGCATGTTTGGCAGTG 1080
QY 4859 TTCTTCTCTAGCAAGAGTGAAGTCAAGGCTTTCAGTGTCTTCAAGTGTCTCGCT 4918
DB 1081 TTCTTCTCTAGCAAGAGTGAAGTCAAGGCTTTCAGTGTCTTCAAGTGTCTCGCT 1140
QY 4919 GTTCTTGTGATTAACAAGCTGCGATCTTCGCTGTGAACCTTCGCTGCGCTGGCCGCTC 4978
DB 1141 GTTCTTGTGATTAACAAGCTGCGATCTTCGCTGTGAACCTTCGCTGCGCTGGCCGCTC 1200
QY 4979 TACATTTGATCGATCGATCCGATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 5038
DB 1201 TACATTTGATCGATCGATCCGATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 1260
QY 5039 GCAATTTGCTGGGCGCTGCAAGTCTTTCGCTGTGATTTGATGATGCTCGATGATG 5098
DB 1261 GCAATTTGCTGGGCGCTGCAAGTCTTTCGCTGTGATTTGATGATGCTCGATGATG 1320
QY 5099 GTTCATGTTCTTGGTGTGGCGGTGATGATCTTGTATGATGATGATGATGATGATGATGAT 5158
DB 1321 GTTCATGTTCTTGGTGTGGCGGTGATGATCTTGTATGATGATGATGATGATGATGATGAT 1380
QY 5159 TATGGCTTTACTGTTGCTGCGCAACGCGATGATGATGATGATGATGATGATGATGATGAT 5218
DB 1381 TATGGCTTTACTGTTGCTGCGCAACGCGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 5219 GTGCTTCAGAGAACGACCAAGCCAGAGAGAGACCCGAGAAATTTTCAACGATTC 5278
DB 1441 GTGCTTCAGAGAACGACCAAGCCAGAGAGAGACCCGAGAAATTTTCAACGATTC 1500
QY 5279 ACCATCATCCAGGACCTTTGACCGGTGAAGCTATGCACTGAGAGCGTCAAGCATGCC 5338
DB 1501 ACCATCATCCAGGACCTTTGACCGGTGAAGCTATGCACTGAGAGCGTCAAGCATGCC 1560
QY 5339 ATGTTTGCAGCGGAAAGCTTGGCTGAGTGTGATGATGATGATGATGATGATGATGATGAT 5398
DB 1561 ATGTTTGCAGCGGAAAGCTTGGCTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1620
QY 5399 GTTTCACAGTGAAGCGGAAAGATCGTGTGCTTCCATCTGTCACCTTTCCGAGTC 5458
DB 1621 GTTTCACAGTGAAGCGGAAAGATCGTGTGCTTCCATCTGTCACCTTTCCGAGTC 1680
QY 5459 CGCACTAAGGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5518
DB 1681 CGCACTAAGGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 5519 GTTAACTCAAGCGGACGATTTAAGCGCTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 5578
DB 1741 GTTAACTCAAGCGGACGATTTAAGCGCTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
QY 5579 GGGGAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5638
DB 1801 GGGGAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860

```

```

QY 5639 ACCCGATGTTGTTGTTGCAATTACAGAAACCCGACCTGTAAACATTCAGTTTGGGC 5698
DB 1861 ACCCGATGTTGTTGTTGCAATTACAGAAACCCGACCTGTAAACATTCAGTTTGGGC 1920
QY 5699 GAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCAACAA 5758
DB 1921 GAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCAACAA 1980
QY 5759 CCA 5761
DB 1981 CCA 1983

RESULT 5
AAF87497
ID AAF87497 standard; DNA, 1656 BP.
XX
XX AAF87497;
AC
XX
XX 09-JUL-2001 (first entry)
DT
XX
DE Corynebacterium thermaminogenes scrb nucleotide sequence.
XX
KW Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW thermolactant; acea; accBC; dtsR; dtsR2; pfk; scrb; gluaBCD; pdh; pc;
KW ppc; acn; icd; lpd; odh; ds.
XX
OS Corynebacterium thermaminogenes.
XX
FH Key Location/Qualifiers
FT CDS 309..1598
FT /tag= a
FT /product= "scrb protein"
XX
XX WO200125447-A1.
XX
XX 12-APR-2001.
PD
XX
XX 04-OCT-2000; 2000MO-JP006913.
PF
XX
XX 04-OCT-1999; 99JP-00282716.
PR 01-NOV-1999; 99JP-00311147.
PR 21-APR-2000; 2000JP-00120687.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K;
XX Kimura E, Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
XX Sugimoto S;
XX WPI; 2001-300170/31.
XX P-PSDB; AAB83190.
XX
XX Proteins and their DNA useful for microbial production of L-amino acids.
XX
XX Claim 27, Page 199-201, 215pp; Japanese.
XX
XX The present sequence is provided in a specification relating to genes
XX encoding thermophilic amino acid biosynthesis system enzymes of the
XX thermolactant bacterium Corynebacterium thermaminogenes. The novel
XX proteins retain at least 30% isocitrate lyase activity after heating at
XX 500C for 5 minutes. DNA fragments encoding the enzymes were isolated from
XX a Corynebacterium thermaminogenes chromosomal DNA plasmid library by
XX PCR. The DNA may be used for developing strains of amino acid producing
XX microorganisms
XX
SQ Sequence 1656 BP; 392 A; 456 C; 432 G; 376 T; 0 U; 0 Other;

Query Match 24.4%; Score 1457; DB 5; Length 1656;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1545; Conservative 0; Mismatches 110; Indels 2; Gaps 2;

QY 2031 ACGCTGATTTCTTCAACACCATCGAAGAGTCCCAACCAAGCGCTCAACCAAGGTTGG 2090

```



```
Db 1 |ACGCCGATTTCTTCAACTATCGAAGAGGTCCCAACCCAGCTTGACCCAGGCGTTGG 60
Qy 2091 GCACCTTTGTCGCCGCGCGCAAAACATCGTGTGTGTGGCAACTGTGTGAAGAAAAGCCGAC- 2149
Db 61 GTACTTTTGTCCGCGCGCAAAATATCGTGTGTGTGTGGCAACTGTGTGCAAGAAAAGCAGACA 120
Qy 2150 GCCATTCGCGGAACTGTGTGAAGGCCCACTGTACTGCTTCTTGCCGAGTTTCCATCTGTAG 2209
Db 121 GCCATTCGCGGAACTGTGTGAAGGTCCAGTACTGCTTCTTGCCGAGTTTCCATCTGTGA 180
Qy 2210 ATGCAACAATGCAACCATCATCTGTGTGAAGAGCAGATATCCAAAGTGTGAAAACGT 2269
Db 181 ATGCAACAACAGCCACCATCATCTGTGTGAAGAGCAGATATCCAAAGTGTGAAAATGCT 240
Qy 2270 GATCACTACCGTCTCATGTGAGCAATTAAGCTGCGCTGAGAAAACAAAAGAAAGTACTGT 2329
Db 241 GACCACTTACGTCCTCATGTGAGCAATTAAGCTGCGCTGAGAAAACAAAAGAAAGTACTGT 300
Qy 2330 GTGGGGCTATGACACACAACTTTTCAAGTTTGGCCCTGCGTACCATGTGATCTCTCCGC 2389
Db 301 GTGGGGCTATGACACACAACTTTTCAAGTTTGGCCCTGCGTACCATGTGATCTCTCCGC 360
Qy 2390 AGGCGAGGCTCAATGATCCCAACGGAATGTACGTGATGAGATACCTCCACGTCCTACT 2449
Db 361 AGGCGAGACTCAATGATCCCAATGGAATGTACGTGATGAGATACCTCCACGTCCTACT 420
Qy 2450 ACCAGCAGATCCAGGTTTCCCTTTGCAACCAAGCGCACCGGCTGGGCTCACACACCA 2509
Db 421 ACCAGCAGATCCAGGTTTCCCTTTGCAACCAAGCGCACCGGTTGGGCTCACACACCA 480
Qy 2510 CGCCGTGACCGGACCGGACATGTGAGTGAACCACTGCGCGACGCTCTTACCCGG 2569
Db 481 CGCCGTGACCGGACCGGACATGTGAGTGAACCACTGCGCGACGCTCTTACCCGG 540
Qy 2570 ATGCATCTCATGACCTGATGATGATCTATTCGCTGAGAGCCGATTTTACTGACGCGACAC 2629
Db 541 ATGCATCTCATGACCTGATGATGATCTATTCGCGGAGAGCCGATTTTCTGACGCGACAC 600
Qy 2630 TTTAACTTTTCTACACCGGCAACTTAAATTTGACGGAAGGCGCGCCACCCAAAAC 2689
Db 601 TTTAACTTTTCTACACCGGCAACCGAAAATTTGACGGAAGGCGCGCCACCCAAAAC 660
Qy 2690 TTGTGAAGTGAAGACCCCACTGGGCTGATGTGGGCGCAATTCATGCGGCTGCTTAA 2749
Db 661 TTGTGAAGTGAAGACCCCACTGGGCTGATGTGGGCGCAATTCATGCGGCTGCTTAA 720
Qy 2750 ATCCGCTTATCGACGGAACCGCGACGCGGTTTCAACCCCATTTACCGCATCCCATGATCA 2809
Db 721 ATCCGCTTATCGACGGAACCGCGACGCGGTTTACGCCCACTACCGCATCCCATGATCA 780
Qy 2810 GCCCTGATGTGATGTGTGAACAATGTTCTTGCGGCGCCCAACGCGAAAACCTCACCGGTG 2869
Db 781 GCCCTGATGTGATGTGTGAAGAATGTTCTTGCGGCGCTCACGCGAAAACCTCACCGGTG 840
Qy 2870 CAGCGGTTCTATACCGCTCGACAGATCTTGAACCTGGGAATTTCTCCGCTGAATACACT 2929
Db 841 CAGCGGTTCTATACCGCTCGACAGATCTTGAACCTGGGAATTTCTCCGCTGAATACACT 900
Qy 2930 TTGACCTCAGTATGACAACTGTGTTCTGCTCTGATCTCGTTCCGATGCTACATGT 2989
Db 901 TTGACCTCAGTATGACAACTGTGTTCTGCTCTGATCTCGTTCCGATGCTACATGT 960
Qy 2990 GGGAAATGCCCAACTTTTACGCTTCCGCGATGAAGAAAATGCGGAGATCTCGACGTCG 3049
Db 961 GGGAAATGCCCAACTTTTACGCTTCCGCGATGAAGAAAATGCGGAGATCTCGACGTCG 1020
Qy 3050 TGAATTTCTGTCACAAAGATTTGAACCGAATTCACAGATGAGTTACTACTAGCAAGT 3109
Db 1021 TGAATTTCTGTCACAAAGATTTGAACCGAATTCAGATGATGAGTTACTACTAGCAAGT 1080
Qy 3110 CTGACAGTGCAGATATGTCTGCAACAAGCTTGAAGAAAGCACTTTCGCGCTTTCGAG 3169
```

```
Db 1081 CTGACAGTGGGATATGTCTGCGCAAGCTTGAAGAAAGCACTTCGCTGCTGGAG 1140
Qy 3170 GATTACGAGACTGATTTGCGCCATGATATTTACGACCGGAGTTGCACTAAACGTT 3229
Db 1141 GATTACGAGACTGATTTGCGCTATGATATTTACGCGCCGAGTTGCACTAAACGTT 1200
Qy 3230 CTGATGCTGCTGCTGCTGAGATGGGCGCTGCGCGAGAGATGATCACCAACGTTG 3289
Db 1201 CGATGCTGCTGCTGCTGAGATGGGATGCTTGCACAGAGATGATCACCAACGTTG 1260
Qy 3290 CACAGAAAGATGAGTGCATGCTGCTGACCTGCTGAGCCCGCAAGCTTCAATTTGGCAACG 3349
Db 1261 CGAGAAAGATGAGTGTGACCTGCTGACCTGCTGACCTGCTGACCTGCTGCTGCTGCTG 1320
Qy 3350 CGATCTACCAAGACTCTCTTCTCCAGAGGGGAGTGGGGGCTAATCAGATCTGTATTAG 3409
Db 1321 CGATCTACCAAGACTCTCTTCTCCAGAGGGGAGTGGGGGCTAATCAGATCTGTATTAG 1380
Qy 3410 GTTCTGAACCTGCTCGAGTACATCCGAGGCAATATTTCCTGAGTGGGATGGTGC 3469
Db 1381 GTTCTGAACCTGCTCGAGTACATCCGAGGCAATATTTCCTGAGTGGGATGGTGC 1440
Qy 3470 GTTGTCTGTGATCTGTATGATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3529
Db 1441 GTTGTCTGTGATCTGTATGATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Qy 3530 TGAATGCGGAGATATACACCATTTGAGATTAATGCAAGTGTGATTTTCAATTCG 3589
Db 1501 TGAATGCGGAGATATACACCATTTGAGATTAATGCAAGTGTGATTTTCAATTCG 1560
Qy 3590 CTTTTCGCGGCTTCAAAAGTGAACACTATTTAGAGATTAATGATTAAGGATCTTTG 3649
Db 1561 C-TTTCGCGGCTTCAAAAGTGAACACTATTTAGAGATTAATGATTAAGGATCTTTG 1619
Qy 3650 TGGCGAATTTGACAAATCTTGCAGAAATCCCTTGAT 3686
Db 1620 TGGCGAATTTGACAAATCTTGCAGAAATCCCTTGAT 1656
```

```
RESULT 6
AAFP1528
ID AAFP1528 standard; DNA; 1527 BP.
AC AAFP1528;
XX
XX
XX 09-APR-2001 (first entry)
DT
XX
XX
XX C-glutamicum phosphoenolpyruvate DNA #1.
DE
XX
XX Phosphoenolpyruvate; sugar phosphotransferase system; PHS; ds.
KW
XX
XX Corynebacterium glutamicum.
OS
XX
XX WO200102583-A2.
PN
XX
XX 11-JAN-2001.
PD
XX
XX 27-JUN-2000; 2000WO-IB000973.
PF
XX
XX 01-JUL-1999; 99US-0142691P.
PR 23-AUG-1999; 99US-0150310P.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042097.
XX
XX (BADI ) BASF AG.
XX
XX Pompejus M, Kroege B, Schroeder H, Zelder O, Habermann G;
PI WPI; 2001-080989/09.
XX
XX Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:
PT sugar phosphotransferase system proteins or their portions, useful for
XX typing or identifying C. glutamicum or related bacteria, and as markers
PT
```

PT for transformation.

XX PS Claim 3; Page 98-101; 144bp; English.

CC The present invention relates to Corynebacterium glutamicum
CC phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The
CC PTS nucleic acids and proteins are useful in the identification of
CC microorganisms which can be used to produce fine chemicals, for
CC modulating fine chemical production in C. glutamicum or related bacteria,
CC the typing or identification of C. glutamicum or related bacteria, as
CC reference points for mapping C. glutamicum genome, and as markers for
CC transformation

XX Sequence 1527 BP; 304 A; 392 C; 430 G; 401 T; 0 U; 0 Other;

Query Match 23.7%; Score 1412.6; DB 5; Length 1527;
Best Local Similarity 95.7%; Pired. No. 0;
Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

```

QY 4259 CTGATGGCACTGCGCGGTTGCGGTTCTTGCCAGTGTGGTTTTCACCCGAACCAAG 4318
DB 1 CTCATGGCACTGCGCGGTTGCGGTTCTTGCCAGTGTGGTTTTCACCCGAACCAAG 60
QY 4319 CGTTTCGATGCGCAATGATGTTCTGCGCGCGG-GCATTGGTATGGCGATGTTCCCAAC 4377
DB 61 CGTTTCGCGGCAATGATGTTCTGCGCGCGGCGGATTTGGTATGGCGATGTTCCCGAG 120
QY 4378 CCTGGTAAACGGCTACGACGTGGCGCCACACATGACCGCGGCGAAATGCCAATGTGTTC 4437
DB 121 CTGTGGAAACGGCTACGACGTGGCGCCACACATGCGTGGCGGCGAAATGCCAATGTGTTC 180
QY 4438 CCGTTGGTGGATGGATGTTGCTCAAGCTGGTTACGAGGCGACCGTCTCTGTGTGTGT 4497
DB 181 CCGTTGGTGGATGGATGTTGCTCAAGCTGGTTACGAGGCGACCGTCTCTGTGTGTGT 240
QY 4498 GGTCTCTGGATCTGCGCAACGATCGAAGATTTCTGCAACAAGCACTATGGGCACTGC 4557
DB 241 GGTCTCTGGATCTGCGCAACGATCGAAGATTTCTGCAACAAGCACTATGGGCACTGC 300
QY 4558 AAGACTTCTGATCAACCCAGTGTGTAAGTCTGCTGCTACCGCGTCTCTTACGTTCAATTC 4617
DB 301 AAGACTTCTGATCAACCCAGTGTGTAAGTCTGCTGCTACCGCGTCTCTTACGTTCAATTC 360
QY 4618 TATTTGGTCCAGAAATGCGTGGTGGTGAATCTTGTCTGCGACAACGCTGTGAGAGATCTTA 4677
DB 361 CATTTGCCAGAAATGCGTGGTGGTGAATCTTGTCTGCGACAACGCTGTGAGAGATCTTA 420
QY 4678 TGAATTCGATGCTCAGTGGCGGCTGCTTTTCCGTTCCGTTCTGATCTACCAATCGTAT 4737
DB 421 TGAATTCGATGCTCAGTGGCGGCTGCTTTTCCGTTCCGTTCTGATCTACCAATCGTAT 480
QY 4738 CACTGGTCTGCAACAGTCTCTCCCGCAATTGAGCTGAGCTGTTCACACAGGCTGATC 4797
DB 481 CACTGGTCTGCAACAGTCTCTCCCGCAATTGAGCTGAGCTGTTCACACAGGCTGATC 540
QY 4798 CTTCAATCTTCCGCAACCGCATTCATGCGCAATATCGCGAGGGTGCAGCATGTTTGGCAAT 4857
DB 541 CTTCAATCTTCCGCAACCGCATTCATGCGCAATATCGCGAGGGTGCAGCATGTTTGGCAAT 600
QY 4858 GTTCTTCCAGGAAAGAGTAAAGTCAAGGCGCTTGAGAGGCTTCAAGTGTCTCCGCG 4917
DB 601 GTTCTTCCAGGAAAGAGTAAAGTCAAGGCGCTTGAAGGCTTCAAGTGTCTCCGCG 660
QY 4918 TGTTCCTGGTATTAACAGACCTGCGATCTTGGTGTGAACCTTTCGCTGCGTGGCGCTT 4977
DB 661 TGTTCCTGGTATTAACAGACCTGCGATCTTGGTGTGAACCTTTCGCTGCGTGGCGCTT 720
QY 4978 CTACATTTGGTATCGGATACCGGATATCGGTGGCGCTTGTGATTCATCTTTGATATCA 5037
DB 721 CTTCATTCGATTCGATACCGGATATCGGTGGCGCTTGTGATTCATCTTTGATATCA 780
QY 5038 GGCAGTTGGTGGCGCTGCGATGTTCTTGGGTGTGTTTCTATGATGCTCCAGATAT 5097

```

```

DB 781 GGCAGTTGGTGGCGCTGCGATGTTCTTGGGTGTGTTTCTATGATGCTCCAGATAT 840
QY 5098 GGTATGTTCTTGTGTTGCGCGGTAGTAACTTTGTATGCAATTCGCGCGACGATGTC 5157
DB 841 GGTATGTTCTTGTGTTGCGCGGTAGTAACTTTGTATGCAATTCGCGCGACGATGTC 900
QY 5158 TTATGGCTTTTCTTGTGTTGCGCGAGGCAAGGCAATTTGATTCAGATGCAACCGTCTCC 5217
DB 901 TTATGGCTTTTCTTGTGTTGCGCGAGGCAAGGCAATTTGATTCAGATGCAACCGTCTCC 960
QY 5218 AGTGCCTGAGGAACGACCAAGCCGAAAGCAGAAAGCAGACCCGGAATTTTCAACGATTC 5277
DB 961 AGTGCCTGAGGAACGACCAAGCCGAAAGCAGAAAGCAGACCCGGAATTTTCAACGATTC 1020
QY 5278 CACCATATCAGGCACTTTTGAACCGGTGAGCTATGCACTGAGCAGCTGACGATGC 5337
DB 1021 CACCATATCAGGCACTTTTGAACCGGTGAGCTATGCACTGAGCAGCTGACGATGC 1080
QY 5338 CATGTTTGCAGCGAAAGCTTGGCTCAGGTGTGCGATGTCGCCCAAGGAGGAGCT 5397
DB 1081 CATGTTTGCAGCGAAAGCTTGGCTCAGGTGTGCGATGTCGCCCAAGGAGGAGCT 1140
QY 5398 GGTTCACACGATGAGCGGAAAGATCGTGGTGCCTTCCCATGTGTCAGCTTTCGAGT 5457
DB 1141 AGTTTCTCCGTTGAGTGAAGAAATGTGTGTCATTCATCTGCGCATGCTTTCGAGT 1200
QY 5458 CCGCATTAAGGCTGAGATGTTTCAATGTGATATCTTGTATGCACTTGTGTTTCAGAC 5517
DB 1201 TCGCACAAAGGCTGAGATGTTTCAATGTGATATCTTGTATGCACTTGTGTTTCAGAC 1260
QY 5518 CGTAAACCTCAACGGCAGCAGCACTTTAACCCGCTGAAGAAAGCGGCGAGTCAAAAC 5577
DB 1261 AGTAAACCTCAACGGCAGCAGCACTTTAACCCGCTGAAGAAAGCGGCGAGTCAAAAC 1320
QY 5578 AAGGAGAGCTGCTGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 5637
DB 1321 AAGGAGAGCTGCTGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 5638 CAGCGGATGTTGTTTGAATTAACAAGAAACGGAACCTGTAAACCTTACGTTTGGG 5697
DB 1381 CAGCGGATGTTGTTTGAATTAACAAGAAACGGAACCTGTAAACCTTACGTTTGGG 1440
QY 5698 CGAAATTAAGGAGGAGCAACCTGCTCAACCTGCCAAAGAAAGAGGCTGCCAGAAC 5757
DB 1441 CGAAATTAAGGAGGAGCAACCTGCTCAACCTGCCAAAGAAAGAGGCTGCCAGAAC 1500
QY 5758 ACCATTAAGTGAACCTTGAAGTTCG 5784
DB 1501 ACCATTAAGTGAACCTTGAAGTTCG 1527

```

RESULT 7
AB865346
ID AB865346 strand: DNA; 1527 BP.
XX
AC AB865346;
XX
DT 15-NOV-2002 (first entry)
XX
DE DNA encoding C. glutamicum metabolic pathway (MP) protein #5.
XX
KW Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;
KW cofactor; nucleotide; nucleoside; trehalose; fine chemical production;
KW organic acid; non-proteinogenic amino acid; purine base; carbohydrate;
KW pyrimidine base; lipid; unsaturated fatty acid; diol; polyketide;
KW aromatic compound; food industry; animal feed; cosmetic industry;
KW pharmaceutical industry; gene; de.
OS Corynebacterium glutamicum ATCC 13032.
XX
PN WO200251231-A1.
XX
PD 04-JUL-2002.

XX 22-DEC-2000; 2000WO-EP013143.
 XX 22-DEC-2000; 2000WO-EP013143.
 XX (BADI) BASF AG.
 XX Pompejus M, Kroegeer B, Zelder O, Schroeder H;
 XX MPI; 2002-643289/69.
 DR P-PSDB; ABG80325.
 PT New metabolic pathway genes of *Corynebacterium glutamicum* for producing
 PT fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,
 PT cofactors or enzymes used in food, feed, cosmetics or pharmaceutical
 PT industries.
 PS Claim 1; Page 108-111; 176pp; English.
 XX The present invention relates to the isolation of *Corynebacterium*
 CC glutamicum metabolic pathway (MP) proteins, and the polynucleotide
 CC sequences encoding them. The MP proteins are enzymes involved in the
 CC metabolism of molecules important for the normal functioning of cells
 CC (e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or
 CC trehalose). The polynucleotide sequences encoding the MP proteins are
 CC useful for producing fine chemicals, particularly organic acids, non-
 CC proteinogenic amino acids, purine and pyrimidine bases, nucleosides,
 CC nucleotides, lipids, (un)saturated fatty acids, diols, carbohydrates,
 CC aromatic compounds, vitamins, cofactors, polyketides and enzymes. The
 CC fine chemicals are useful in the food, animal feed, cosmetic or
 CC pharmaceutical industries. ABS65342-ABS65364 encode the C. glutamicum MP
 CC proteins of the invention
 SQ Sequence 1527 BP; 304 A; 392 C; 430 G; 401 T; 0 U; 0 Other;

Query Match 23.7%; Score 1412.6; DB 6; Length 1527;
 Best Local Similarity 95.7%; Pred. No. 0;
 Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

QY 4259 CTGATGCACTGCGCGCTTGGCTTCTGCAAGTGTGTTGTTTCCGCAACCAAG 4318
 DB 1 CTCATGCACTGCGCGCTTGGCTTCTGCAAGTGTGTTTCCGCAACCAAG 60
 QY 4319 CGTTTGGTGGCAATAGTTCCTGGGCGCGG-CGATTGGTATGGCAATGTGTCCCAAC 4377
 DB 61 CGTTTGGGCGCAATAGTTCCTGGGCGCGGCGTATGTTATGGCATGTGTCCCGAG 120
 QY 4378 CCGTGTAAACGGGCTACGAGTGGCGCGCAACATGACCGGCGGGAATGCCAATGTGTC 4437
 DB 121 CTTGTGGAACGGGCTACGAGTGGCGCGCAACATGAGTGGCGGGAATGCCAATGTGTC 180
 QY 4438 CCTGTTGGTGTGATGTGCTCAAGCTGTGTTACAGGGGACCGGTCTCTGTGTGTGT 4497
 DB 181 CCGTTTGGTGTGATGTGCTCAAGCTGTGTTACAGGGGACCGGTCTCTGTGTGTGT 240
 QY 4498 GGTCTTTGGTGTGATGTGCTCAAGCTGTGTTACAGGGGACCGGTCTCTGTGTGTGT 4557
 DB 241 GGTCTTTGGTGTGATGTGCTCAAGCTGTGTTACAGGGGACCGGTCTCTGTGTGTGT 300
 QY 4558 AGACTTCTGATCAACCCAGTGTGCTGCTGCTCAACCGGCTTCTTACGTTCACTTGC 4617
 DB 301 AGACTTCTGATCAACCCAGTGTGCTGCTGCTCAACCGGCTTCTTACGTTCACTTGC 360
 QY 4618 TATTTGTCAGCAATCGCTGGGTGGGTGAATTGCTGGGCAACGGTCTGAGGAGACTTA 4677
 DB 361 CATTTGCCAGCAATCGCTGGGTGGGTGAATTGCTGGGCAACGGTCTGAGGAGACTTA 420
 QY 4678 TGAATTCGGTGGTGGGCGGTGCTGTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4737
 DB 421 TGAATTCGGTGGTGGGCGGTGCTGTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
 QY 4738 CACTGCTGCAACCACTCTTCCCGCAATTGAGCTGAGCTGTTCAACGAGGTGATC 4797

DB 481 CACTGCTGCAACCACTCTTCCCGCAATTGAGCTGAGCTGTTTCAACGAGGTGATC 540
 QY 4798 CTTCACTTTGGCAACCGCATCATGCGCAATATGCGCAGGCTGACAGATTTTGGCAGT 4857
 DB 541 CTTCACTTTGGCAACCGCATCATGCGCAATATGCGCAGGCTGACAGATTTTGGCAGT 600
 QY 4858 GTTCTTCTAGCAAGAGTGAAGAGCTCAAGGGGCTTGCAGGTGCTGCTCGC 4917
 DB 601 GTTCTTCTAGCAAGAGTGAAGAGCTCAAGGGGCTTGCAGGTGCTGCTCGC 660
 QY 4918 TGTCTTGTGATTAACAGAGCTGCGATCTTGTGTGTGAACCTTGCGCTGCGCTG 4977
 DB 661 TGTCTTGTGATTAACAGAGCTGCGATCTTGTGTGTGAACCTTGCGCTGCGCTG 720
 QY 4978 CTACATTTGATTCGGTACCGGAGCTATGCGTGGGCTTGTATGCACTTTGATATCA 5037
 DB 721 CTTCATCGGATTCGGTACCGGAGCTATGCGTGGGCTTGTATGCACTTTGATATCA 780
 QY 5038 GGCAGTTGCGTTGGGCGCTGAGAGTTCTTGGGTGTGTTCTATGATGCTCCAGATAT 5097
 DB 781 GGCAGTTGCGTTGGGCGCTGAGAGTTCTTGGGTGTGTTCTATGATGCTCCAGATAT 840
 QY 5098 GGTCACTGTTCTTGTGTTGGCGGTAATACCTTTGTCAATGCAATTCGCGCAGATTC 5157
 DB 841 GGTCACTGTTCTTGTGTTGGCGGTAATACCTTTGTCAATGCAATTCGCGCAGATTC 900
 QY 5158 TTATGCGCTTTATCTTGTGTTGGCGGTAATACCTTTGTCAATGCAATTCGCGCTCC 5217
 DB 901 TTATGCGCTTTATCTTGTGTTGGCGGTAATACCTTTGTCAATGCAATTCGCGCTCC 960
 QY 5218 AGTGCCTGAGAAACGACCAAGCGGAGAGCAAGAGCAACCGCAGATTTCAAGATTC 5277
 DB 961 AGTGCCTGAGAAACGACCAAGCGGAGAGCAAGAGCAACCGCAGATTTCAAGATTC 1020
 QY 5278 CACCATCATCAAGGACCTTTGACCGGTGAAGCTATGCACTGAGACGCTGACGATTC 5337
 DB 1021 CACCATCATCAAGGACCTTTGACCGGTGAAGCTATGCACTGAGACGCTGACGATTC 1080
 QY 5338 CATGTTTGGCAGGGAAGCTTGGCTCAGGTGTGCGATTCGTTCCCAACCAAGGGGCACT 5397
 DB 1081 CATGTTTGGCAGGGAAGCTTGGCTCAGGTGTGCGATTCGTTCCCAACCAAGGGGCACT 1140
 QY 5398 GGTTCACACAGTGAACGGAAGATGCTGTGGCTTCCATCTGATGCAAGCTTTCGAGT 5457
 DB 1141 AGTTCCTCGGTGATGGAAGATTTGTGTGATTCATCTGCGCAATGCTTTCGAGT 1200
 QY 5458 CCGCATTAAGGCTGAGATGTTCCATGTGATATCTTGAATGCAATTTGTTTCGATC 5517
 DB 1201 CCGCATTAAGGCTGAGATGTTCCATGTGATATCTTGAATGCAATTTGTTTCGATC 1260
 QY 5518 CGTAAACCTCAACGGCAGCACTTTAACCGCTGAAGAGAGGCGCATGAAGTCAAGG 5577
 DB 1261 AGTAAACCTCAACGGCAGCACTTTAACCGCTGAAGAGAGGCGCATGAAGTCAAGG 1320
 QY 5578 AGGGAGCTGCTGTGATATGATATGATGATGATGATGATGATGATGATGATGATGAT 5637
 DB 1321 AGGGAGCTGCTGTGATATGATATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 5638 CACGCGATGTGTTTGTGATTAACAGAAAAACGGACCTGTAAACCTTACGTTTGG 5697
 DB 1381 CACGCGATGTGTTTGTGATTAACAGAAAAACGGACCTGTAAACCTTACGTTTGG 1440
 QY 5698 CGAAATTAAGGGAGGCAACCTGCTCAACGTCGCAAGAAAGAGGGGCGCAGAAC 5757
 DB 1441 CGAAATTAAGGGAGGCAACCTGCTCAACGTCGCAAGAAAGAGGGGCGCAGAAC 1500
 QY 5758 ACCATTAAGTGAACCTTGAAGTTCG 5784
 DB 1501 ACCATTAAGTGAACCTTGAAGTTCG 1527

RESULT 8
 AAF71533

ID AAF71533 standard; DNA; 1342 BP.
XX
AC AAF71533;
XX
XX 30-APR-2001 (first entry)
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:347.
XX
XX Corynebacterium glutamicum; carbon metabolism and energy production;
KM SMP protein; sugar metabolism and oxidative phosphorylation protein;
KM fine chemical production; organic acid; pyrimidine base; nucleoside;
KM nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KM nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KM carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KM diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX
XX Corynebacterium glutamicum.
XX
XX WO200100844-A2.
XX
XX 04-JAN-2001.
PD
XX 23-JUN-2000; 2000WO-IB000943.
PF
XX 25-JUN-1999; 99US-0141031P.
XX 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99US-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX
XX (BADI) BASF AG.
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
PI MPI; 2001-061975/07.
XX P-PSDB; AAB79416.
DR
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
XX carbohydrates or enzymes.
XX
XX Claim 3; Page 627-629; 1246pp; English.
XX
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in
CC to AAB 79633 which are involved in carbon metabolism and energy
CC production. The C. glutamicum SMP gene can be used in vectors (II) for

CC expression in host cells and production or modulation of production of
CC fine chemicals, such as, an organic acid, a proteinogenic or
CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
CC encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to C.
CC glutamicum, identify and localise C. glutamicum sequences of interest, in
CC evolutionary studies, in determining SMP protein regions required for
CC function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH)
XX
XX Sequence 1342 BP; 312 A; 373 C; 359 G; 298 T; 0 U; 0 Other;
SQ
XX
XX Query Match 22.1%; Score 1321.2; DB 4; Length 1342;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 2238 ATGAAGACAGCTATCCAAAGCTGAAAAAGCTGATCATCCGTCTCATGAGCAATTAA 2297
DB 1 ATGAAGACAGCTATCCAAAGCTGAAAAAGCTGATCATCCGTCTCATGAGCAATTAA 60
QY 2298 ACCTGGCGCTAAGAAACAAAAGAAAGTACTGTGGGGCTATGACACAGAACTTCCAG 2357
DB 61 ACCTGGCGCTAAGAAACAAAAGAAAGTACTGTGGGGCTATGACACAGAACTTCCAG 120
QY 2358 TTTGGGCGCTGCTACATGTACTCTCCGAGGGCAGGCTCAATGATCCCAAGGAT 2417
DB 121 TTTGGGCGCTGCTACATGTACTCTCCGAGGGCAGGCTCAATGATCCCAAGGAT 180
QY 2418 GTACGTGATGAGATACCCCTTCACGCTTACTACAGACGATCCAGGTTCCCTTCCG 2477
DB 181 GTACGTGATGAGATACCCCTTCACGCTTACTACAGACGATCCAGGTTCCCTTCCG 240
QY 2478 ACCAAAGCGACCGGCTGGGCTCACACACACGCGCTGACCGGACCGCAAGTTGA 2537
DB 241 ACCAAAGCGACCGGCTGGGCTCACACACACGCGCTGACCGGACCGCAAGTTGA 300
QY 2538 GTGACGACGCTGCGCCGAGCTCTTAAACCGGATCATCTTATGACCTGATGATCTA 2597
DB 301 GTGACGACGCTGCGCCGAGCTCTTAAACCGGATCATCTTATGACCTGATGATCTA 360
QY 2598 TTCCGGTGGAGCGATTTACTAGACGACCACTTAACTTTCTACACCGGCAACTTAA 2657
DB 361 TTCCGGTGGAGCGATTTACTAGACGACCACTTAACTTTCTACACCGGCAACTTAA 420
QY 2658 AATTGACGAAAGCGCGCGGCCACCCAAACCTTGTCAAGTCGAGAACCCAACTGGGCT 2717
DB 421 AATTGACGAAAGCGCGCGGCCACCCAAACCTTGTCAAGTCGAGAACCCAACTGGGCT 480
QY 2718 GATGGGCGGCAATTACGCGCGCTTGGCTTAAATCGCTTATCGACGACCGCGCGG 2777
DB 481 GATGGGCGGCAATTACGCGCGCTTGGCTTAAATCGCTTATCGACGACCGCGCGG 540
QY 2778 TTTCACACCCCATTAACCGGATCCCATGATCAGCCCTGATGATGATGATGATGAT 2837
DB 541 TTTCACACCCCATTAACCGGATCCCATGATCAGCCCTGATGATGATGATGATGAT 600
QY 2838 TCTTGGGGCGCCCAAGCGGAAACCTTCAACCGGTGACGCGTCTTATACCGTGCACGAT 2897
DB 601 TCTTGGGGCGCCCAAGCGGAAACCTTCAACCGGTGACGCGTCTTATACCGTGCACGAT 660
QY 2898 TGAATACTGGGAATTTCTCGGCGTGAATACCTTTGACCTCACTGATGACCACTGGTTC 2957
DB 661 TGAATACTGGGAATTTCTCGGCGTGAATACCTTTGACCTCACTGATGACCACTGGTTC 720
QY 2958 TGCCTCTGATCTGTTCCCGATGAGGTACATGATGAGAAATGCCCAACTTTTACGCTTGG 3017
DB 721 TGCCTCTGATCTGTTCCCGATGAGGTACATGATGAGAAATGCCCAACTTTTACGCTTGG 780

QY	3018	CGATGAAGAAACTGGCCGAAGATCTGCACTGTGATTTTTCGTCCACAAGATTGACCG	3077
Db	781	CGATGAAGAACTGGCCGAAGATCTGCACTGTGATTTTTCGTCCACAAGATTGACCG	840
QY	3078	AATCCACGATGAGGTTACTCACTACGCAAGCTCTGACCAAGTGCAGATATGTCTGACAA	3137
Db	841	AATCCACGATGAGGTTACTCACTACGCAAGCTCTGACCAAGTGCAGATATGTCTGACAA	900
QY	3138	GCTTGAAGGAACGACCTTCCGGCTTTTGGCAGAGATTCAGCCAGCTGGAATTTGGGCATGA	3197
Db	901	GCTTGAAGGAACGACCTTCCGGCTTTTGGCAGAGATTCAGCCAGCTGGAATTTGGGCATGA	960
QY	3198	ATTTCAGCACCGCAGGTTGACAGTAAAGGATCTGATGCTGGCTCGAGGGCTGGATGG	3257
Db	961	ATTTCAGCACCGCAGGTTGACAGTAAAGGATCTGATGCTGGCTCGAGGGCTGGATGG	1022
QY	3258	GCTGCCCGCGCAGATGATCAACCAACAGTTGCAAGGAAGATGGGTGCACCTGCCATAC	3317
Db	1021	GCTGCCCGCGCAGATGATCAACCAACAGTTGCAAGGAAGATGGGTGCACCTGCCATAC	1080
QY	3318	TGTGCCCCGCAAGCTTCATTTCGCCAACCAACGATCTACCAAGAGCTCTTCTCCAGA	3377
Db	1081	TGTGCCCCGCAAGCTTCATTTCGCCAACCAACGATCTACCAAGAGCTCTTCTCCAGA	1140
QY	3378	GGGGGAGTCGGGGGAAATCAGATCTGTATTAAGTCTGAAACCTGTCCAGATAGACATCG	3437
Db	1141	GGGGGAGTCGGGGGAAATCAGATCTGTATTAAGTCTGAAACCTGTCCAGATAGACATCG	1200
QY	3438	AGGCATATTTTCCCTCGAGTGGGATGAGTCCGTTGTCTGTGATCTGATAGTGATCG	3497
Db	1201	AGGCATATTTTCCCTCGAGTGGGATGAGTCCGTTGTCTGTGATCTGATAGTGATCG	1260
QY	3498	TGCGGTAGCTAGATAAACCTGCGGAATTAAGTATCGCGGACGATTAATACACCATTTGA	3557
Db	1261	TGCGGTAGCTAGATAAACCTGCGGAATTAAGTATCGCGGACGATTAATACACCATTTGA	1320
QY	3558	GATTAATGACAGTGAATGACAG	3579
Db	1321	GATTAATGACAGTGAATGACAG	1342

RESULT	9
AAH67870	
ID	AAH67870 standard; DNA; 1299 BP.
XX	
AC	AAH67870;
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	C glutamicum coding sequence fragment SEQ ID NO: 2905.
XX	
KW	Corynebform bacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis; de.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	EP108790-A2.
XX	
PD	20-JUN-2001.
XX	
PF	18-DEC-2000; 2000EP-00127688.
XX	
PR	16-DEC-1999; 99JP-00377484.
PR	07-APR-2000; 2000JP-00159162.
PR	03-AUG-2000; 2000JP-00280988.
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	
PI	Nakagawa S, Mitsuuchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX	
DR	WPI; 2001-376931/40.
DR	P-PSDB; AAG92651.

XX	Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT	mutation point of a gene, measuring expression of a gene, analyzing
PT	expression profile or pattern of a gene and identifying homologous gene.
XX	
XX	
PS	Claim 8; SEQ ID NO 2905; 246bp + Sequence Listing; English.
XX	
CC	The present invention provides a number of nucleotide and protein
CC	sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC	are useful for identifying the mutation point of a gene derived from a
CC	mutant of corynebacterium bacterium, measuring expression amount and analysing
CC	the expression profile or expression pattern of a gene derived from
CC	Corynebacterium bacterium, and identifying a homologue of a gene derived from
CC	corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
CC	acids, nucleic acids, vitamins, saccharides and organic acids,
CC	particularly L-lysine. The present sequence is a nucleic acid described
CC	in the exemplification of the invention. Note: The sequence data for this
CC	patent did not form part of the printed specification, but was obtained
CC	in electronic format directly from the European Patent Office
XX	
XX	
SEQ	Sequence 1299 BP; 287 A; 363 C; 350 G; 299 T; 0 U; 0 Other;
Query_Match:	21.2%; Score 1267.8; DB 5; Length 1299;
Beet_Local_Similarity	99.0%; Pred. No. 0;
Matches 1286; Conservative	0; Mismatches 12; Indels 1; Gaps 11;

Query Match	Similarity	21.2%	Score 1267.8	DB 5	Length 1299
Best Local	Similarity	99.0%	Pred. No. 0		
Matches 1286	Conservative	0	Mismatches 12	Indels 1	Gaps 1
QY	2337	TGTGTGGGGCTATGACACAGAACTTTTCAGTTTGCGCCCTGCGTACATGTGACTCTC	2386		
DB	2	TGTGTGGGGCTATGACACAGAACTTTTCAGTTTGCGCCCTGCGTACATGTGACTCTC	61		
QY	2387	CGCAGGGCAGGGCTCAATGATATCCCAACGGAAATGTAAGTGAATGAAATCCCTCAACGCT	2446		
DB	62	CGCAGGGCAGGGCTCAATGATATCCCAACGGAAATGTAAGTGAATGAAATCCCTCAACGCT	121		
QY	2447	ACTACACAGACAGATCCAGGTTTCCCTTCGACCCAAACGGACCGGCTCGGCTCACACCA	2507		
DB	122	ACTACACAGACAGATCCAGGTTTCCCTTCGACCCAAACGGACCGGCTCGGCTCACACCA	181		
QY	2507	CCACGCCCTTGACCGGACCCGACGGATTTGACGTGACCGACCTGCGGAGCTCTTTTAC	2566		
DB	182	CCACGCCCTTGACCGGACCCGACGGATTTGACGTGACCGACCTGCGGAGCTCTTTTAC	241		
QY	2567	CGGATGCATCTGTATGACCTGTGATGATGCTATTCGGTGTGAGCCGTATTTACTGACGGCA	2622		
DB	242	CGGATGCATCTGTATGACCTGTGATGATGCTATTCGGTGTGAGCCGTATTTACTGACGGCA	301		
QY	2627	CACCTTAACTTTTCTACACCGGCAACCTTAAATTTGAGAAAGGGCCGCGCACCCAAA	2686		
DB	302	CACCTTAACTTTTCTACACCGGCAACCTTAAATTTGAGAAAGGGCCGCGCACCCAAA	361		
QY	2687	ACCTTGTGGAAGTCGAGGACCCAACTGCGGCTGATGAGGGCGGATTCATGCGCGGTGACCTA	2746		
DB	362	ACCTTGTGGAAGTCGAGGACCCAACTGCGGCTGATGAGGGCGGATTCATGCGCGGTGACCTA	421		
QY	2747	AAATCCGCTTATGACGGACCCCGCAGCGGTTTCAACCCCATTTACCGCGATCCCATG	2804		
DB	422	AAATCCGCTTATGACGGACCCCGCAGCGGTTTCAACCCCATTTACCGCGATCCCATG	481		
QY	2807	TCAGGCCGTGATGATGATGTTGGAACATGATCTTTGGGGGCCCAACGGAAAACCTCACG	2866		
DB	482	TCAGGCCGTGATGATGATGTTGGAACATGATCTTTGGGGGCCCAACGGAAAACCTCACG	541		
QY	2867	GTGCAGCGGTTCTATACCGCTCGACAGATCTTGGAAAACCTGCGGTAATCA	2926		
DB	542	GTGCAGCGGTTCTATACCGCTCGACAGATCTTGGAAAACCTGCGGTAATCA	601		
QY	2927	CCTTTGACCTCAGTATGACCAACCTGATCTGCTCTGATCTGTTTCCGATGGCTACA	2986		
DB	602	CCTTTGACCTCAGTATGACCAACCTGATCTGCTCTGATCTGTTTCCGATGGCTACA	661		
QY	2987	TGTGGGAATGCCCCAACCTTTTTCAGCTTCGCGATGAAGAAACCTGCGAATCTCGAG	3046		
DB	662	TGTGGGAATGCCCCAACCTTTTTCAGCTTCGCGATGAAGAAACCTGCGAATCTCGAG	721		

OY	3047	TGCGATTTTCGTCACAAGGATTTGACCGAATTCACAGATGAGTTACTACATACGAA	3106
Db	722	TGCTGATTTTCCTGTCACAAGGATTTGACCGAATTCACAGATGAGTTACTACATACGAA	781
OY	3107	GCTCTGACCAAGTGGCGGATNTGCTGCACAAGCTTGAAGGAAACGACTTCGCGCTTTGC	3166
Db	782	GCTCGACCAATGGCGGATNTGCTGCGCAAGCTTGAAGGAAACGACTTCGCGCTTTGC	841
OY	3167	GAGGATTCAGCGAGCTGGAATTTCCGCCATGAAATTCACGACCCGAGGTTGCAGTAAACG	3226
Db	842	GAGGATTCAGCGAGCTGGAATTTCCGCCATGAAATTCACGACCCGAGGTTGCAGTAAACG	901
OY	3227	GTTTCGATGCCCTGGCTCGTGGGCTGAGAGGGGGCTCCCGCGCAGAGATGATACCCCAACG	3286
Db	902	GTTTCGATGCCCTGGCTCGTGGGCTGAGAGGGGGCTCCCGCGCAGAGATGATACCCCAACG	961
OY	3287	TTGCAACAGGAAGATGAGTGACCTGCTGACCTGTCGCCCGCAAGCTTATTTGGCAAC	3346
Db	962	TTGCAACGGGAAGATGGGTGCACTGCTGACCTGTCGCCCGCAAGCTTATTTGGCGAAC	1021
OY	3347	ACGCGATTCACAGAGCTCCTTCTCCCAAGGGGGAGTCGGGGGTAAATCAGATCTGTAT	3406
Db	1022	ACGCGATTCACAGAGCTTCTTCTCCCAAGGGGGAGTCAGGGGTAAATCAGATCTGTAT	1081
OY	3407	TAGGTTTCGAACCTGTCGGAATTAACAATCCAGAGCAAAATTTCCCTCGAGTGGGATGATG	3466
Db	1082	TAGGTTTCGAACCTGTCCGAGTGAACATCCAGAGCAAAATTTCCCTCGAGTGGGATGATG	1141
OY	3467	TCCGTTTCTCTGTGGATCGTATGGTGAATCGTCGCGTGTGAGGTAAAACTTGAGCAAT	3526
Db	1142	TCCGTTTCTCTGTGGATCGTATGGTGAATCGTCGCGTGTGAGGTAAAACTTGAGCAAT	1201
OY	3527	TAGTGATCGCGACGATATAACAGCAATGAGATTAATCGCAGGTGATGACAGGTTTCAT	3586
Db	1202	TAGTGATCGCGACGATATAACAGCAATGAGATTAATCGCAGGTGATGACAGGTTTCAT	1261
OY	3587	TCCGTTTTCGCGGCTTCMAAGGTGACACTATTGAGAGA	3625
Db	1262	TCCGCTTCGCGGCTTCMAAGGTGACACTATTGAGAGA	1299

RESULT 10	
AA#71567	
ID	AA#71567 standard; DNA; 1287 BP.
XX	
XX	AA#71567;
XX	
DT	30-APR-2001 (first entry)
XX	
DE	Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:415.
XX	
KW	Corynebacterium glutamicum; carbon metabolism and energy production;
KW	SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW	fine chemical production; organic acid; proteinogenic amino acid;
KW	nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW	nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW	carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW	diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX	
XX	Corynebacterium glutamicum.
OS	
XX	WO200100844-A2.
PN	
XX	
XX	04-JAN-2001.
PD	
XX	
PF	23-JUN-2000; 2000WO-IB000943.
XX	
XX	
PR	25-JUN-1999; 99US-0141031P.
PR	08-JUL-1999; 99DE-01031412.
PR	08-JUL-1999; 99DE-01031413.
PR	08-JUL-1999; 99DE-01031419.
PR	08-JUL-1999; 99DE-01031420.

PR	08-JUL-1999;	99DE-01031428.
PR	08-JUL-1999;	99DE-01031428.
PR	08-JUL-1999;	99DE-01031431.
PR	08-JUL-1999;	99DE-01031433.
PR	08-JUL-1999;	99DE-01031434.
PR	08-JUL-1999;	99DE-01031510.
PR	08-JUL-1999;	99DE-01031562.
PR	08-JUL-1999;	99DE-01031634.
PR	09-JUL-1999;	99DE-01032180.
PR	09-JUL-1999;	99DE-01032227.
PR	09-JUL-1999;	99DE-01032230.
PR	09-JUL-1999;	99US-01432088.
PR	14-JUL-1999;	99DE-01032924.
PR	14-JUL-1999;	99DE-01032973.
PR	14-JUL-1999;	99DE-01033005.
PR	27-AUG-1999;	99DE-01040765.
PR	31-AUG-1999;	99US-01515723.
PR	03-SEP-1999;	99DE-01042076.
PR	03-SEP-1999;	99DE-01042079.
PR	03-SEP-1999;	99DE-01042086.
PR	03-SEP-1999;	99DE-01042087.
PR	03-SEP-1999;	99DE-01042088.
PR	03-SEP-1999;	99DE-01042099.
PR	03-SEP-1999;	99DE-01042123.
PR	03-SEP-1999;	99DE-01042125.

PA (BADI) BASF AG.
XX
XX
PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX
XX WPI: 2001-061975/07.
DR P-PSDB; AAB79450.
DR
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
PT carbohydrates or enzymes.

Claim 3; Page 739-740; 1246pp; English.

CC AAF71361 o AAF71150 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB792433
CC to AAB 79633 which are involved in carbon metabolism and energy
CC production. The C. glutamicum SMP gene can be used in vectors (II) for
CC expression in host cells and production or modulation of production of
CC fine chemicals, such as, an organic acid, a proteogenic or
CC nonproteogenic amino acid (preferred), a purine or pyrimidine base, a
CC nucleoside, a nucleoside, a lipid, a saturated or unsaturated fatty acid,
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
CC polyelectrolyte, or an enzyme. The presence of (I) or SMP proteins (III)
CC encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to C.
CC glutamicum, identify and localise C. glutamicum regions of interest,
CC evolutionary studies, in determining C. glutamicum sequences required for
CC function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH)
CC
CC
CC Sequence 1287 BP; 325 A; 349 C; 326 G; 287 T; 0 U; 0 Other;

Query Match	20.7%;	Score 1237.4;	DB 4;	Length 1287;
Best Local Similarity	97.6%;	Pred. No. 0;		
Matches 1256;	Conservative 0;	Mismatches 31;	Indels 0;	Gaps 0

QY	242	AA	CAGGAGG	CC	CAAGTCCG	GAAGAT	TTAA	CT	TGATCCG	TACATTA	GA	CATCAT	TAC	30.1
Db	1	AA	CAGAGG	CC	CTCAAGTCCG	GAAGAT	TTAA	CT	TGATCCG	TACATTA	GA	CATCAT	TAC	60
QY	302	GT	CTTATG	CTTCTG	TGAGGAA	CCAA	TATAC	CTTCA	GAAAGT	GCAGAG	TGTG	TCATTA	36.1	
Db	61	GT	CTTATG	CTTCTG	TGAGGAA	CCAA	TATAC	CTTCA	GAAAGT	GCAGAG	TGTG	TCATTA	120	

```

QY 362 TCAAGAAATGAGGTCAAGCACTTAATAAATGAGGAGAGATTTGTTCCCTCCCTGG 421
DB 121 TAAAGAAATGAGGTCAAGCACTTAATAAATGAGGAGAGATTTGTTCCCTCCCTGG 180
QY 422 GGTGATGATGAGCTTTCTCAACTGAAAAAGGACATCATCGAAGACTCTGAGAAAC 481
DB 181 GGTGATGATGAGCTTTCTCAACTGAAAAAGGACATCATCGAAGACTCTGAGAAAC 240
QY 482 AGCACTTAATAAAGCAGAGATTCACCCGAACTCCCAAGATTTGTTCCCTGTTTATGA 541
DB 241 AGCACTTAATAAAGCAGAGATTCACCCGAACTCCCAAGATTTGTTCCCTGTTTATGA 300
QY 542 TCTTATATATACGGGTGAGAAAGGTGGCGCTTCTTACGGGAGCGAGACAGCGAG 601
DB 301 TCTTATATATACGGGTGAGAAAGGTGGCGCTTCTTACGGGAGCGAGACAGCGAG 360
QY 602 GAACACCGCGCAGATTCACCCGAACTGAGCAGACCGTATGTTGCCAAGCATGTTTC 661
DB 361 GAATGCCCGCAGTATTCACCCGAACTGAGCAGACCGTATGTTGCCAAGCATGTTTC 420
QY 662 GCGCGCGGTGACGCACTGGCAGCGAGGTGAAAACTTATTCCTGTTGTAAGAGT 721
DB 421 GCGCGCGGTGACGCACTGGCAGCGAGGTGAAAACTTATTCCTGTTGTAAGAGT 480
QY 722 CCTGCTGTGCGGCAATTCACCTGAGGAGCGCTTCTTATCAACGATGCGGTGAGTCA 781
DB 481 CCTGCTGTGCGGCAATTCACCTGAGGAGCGCTTCTTATCAACGATGCGGTGAGTCA 540
QY 782 AAACCCGAGATTCATTTTCCCGGACCCCAAGATCTTCCCGGTATCCATGCGAG 841
DB 541 AAACCCGAGATTCATTTTCCCGGACCCCAAGATCTTCCCGGTATCCATGCGAG 600
QY 842 AAAAGTTGATCAATGATCATCAGTAGCGCGGAACTGACATCTTTCTGACTTCT 901
DB 601 AAAAGTTGATCAATGATCATCAGTAGCGCGGAACTGACATCTTTCTGACTTCT 660
QY 902 CGATCTCTGCGCAGGACCAATCATATGCTCTTCCGGGACACATGATGAGATTTGA 961
DB 661 CGATCTCTGCGCAGGACCAATCATATGCTCTTCCGGGACACATGATGAGATTTGA 720
QY 962 TACCACTACGAGCGAATTTGCTTGGCTAAGAGAAAAATGTGACGGTCACGGCTACGA 1021
DB 721 TACCACTACGAGCGAATTTGCTTGGCTAAGAGAAAAATGTGACGGTCACGGCTACGA 780
QY 1022 TTTGTTCAATGAGATGCTCCGCTCGATCATATGAGGCTCCCGGACCGGTGGCTTGT 1081
DB 781 TTTGTTCAATGAGATGCTCCGCTCGATCATATGAGGCTCCCGGACCGGTGGCTTGT 840
QY 1082 TGCTCGGCACTGCGCGGAGACGATATGTTGATTCGCGAGCGGCTGATTTGGC 1141
DB 841 TGCTCGGCACTGCGCGGAGACGATATGTTGATTCGCGAGCGGCTGATTTGGC 900
QY 1142 CGATGGAACGGTGCATCTAGCTGTTCCAACACCTTTTTCATCAGGACGCGATGGA 1201
DB 901 CGATGGAACGGTGCATCTAGCTGTTCCAACACCTTTTTCATCAGGACGCGATGGA 960
QY 1202 AGCGCGCGGAATGCGAGCGGTGATACATTTTGGGCGTTTGAACGTCAACCTTCACGA 1261
DB 961 AGCGCGCGGAATGCGAGCGGTGATACATTTTGGGCGTTTGAACGTCAACCTTCACGA 1020
QY 1262 TGAAGTCCCGTCTGCGCGATGCGGCGCATTCGCGGAGGACACAGACACTAGCGAG 1321
DB 1021 TGAAGTCCCGTCTGCGCGATGCGGCGCATTCGCGGAGGATACAGACACTAGCGAG 1080
QY 1322 TCAAGTTGTCACACGTCGCGAGGGGTATGACGTTTATGACGCGACCTTCACACTTC 1381
DB 1081 TCAAGTTGTCACACGTCGCGAGGGGTATGACGTTTATGACGCGACCTTCACACTTC 1140
QY 1382 AACCGTCCCGCTAAATTCGTGCTGCGGATACGAATTCGTTAATCCAAACCTTCG 1441
DB 1141 AACCGTCCCGCTAAATTCGTGCTGCGGATACGAATTCGTTAATCCAAACCTTCG 1200
QY 1442 AAATTTGTGCTTTGACTCAACGCGCAGGTGCAAAAGGTCATTTAGTTCATCAAGT 1501

```

```

DB 1201 AAATTTGTGCTTTGACTCAACGCGCACTTACACAGGTCCATTTAGACATCAAGT 1260
QY 1502 ACTTAACTAGAGTAATAACTATCTGT 1528
DB 1261 AATTTAATACGAGCAAACTTCTGT 1287

RESULT 11
AAF71566
ID AAF71566 standard; DNA; 1287 BP.
XX
AC AAF71566;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:413.
XX
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
KW diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
PN W0200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000W0-IB000943.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99DE-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX
PA (BAD) BASF AG.
XX
PI Pompejus M, Kroegeer B, Schroeder H, Zeider O, Habernauer G;
XX WPI; 2001-061975/07.
XX DR P-PSDB; AAB79449.
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar

```


PT typing or identifying C. glutamicum or related bacteria, and as markers
PT for transformation.
XX Claim 5; Page 103-104; 144pp; English.
XX The present invention relates to Corynebacterium glutamicum
CC phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The
CC PTS nucleic acids and proteins are useful in the identification of
CC microorganisms which can be used to produce fine chemicals, for
CC modulating fine chemical production in C. glutamicum or related bacteria,
CC the typing or identification of C. glutamicum or related bacteria, as
CC reference points for mapping C. glutamicum genome, and as markers for
CC transformation
XX
SQ Sequence 1109 BP; 236 A; 278 C; 303 G; 292 T; 0 U; 0 Other;
Query Match 17.6%; Score 1049.8; DB 5; Length 1109;
Best Local Similarity 96.7%; Pred. No. 1.7e-305;
Matches 1072; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 4676 TATGATTTGGTGTGTCAGTGGCGGTCTTTTGGTGTGTTACTTACCAATCGTT 4735
DB 1 TATGATTTGGCGGTGTCAGTGGCGGTCTTTGCTGTGTCTACTACCAATCGTC 60
QY 4736 ATCAGTGTCTGCAAGTCCCTCCCGCAATTGAGCTGTGCAAGCTTCAACGAGGTGA 4795
DB 61 ATCAGTGTCTGCAAGTCCCTCCCGCAATTGAGCTGTGTTAAACGAGGTGA 120
QY 4796 TCCCTCATCTTGCACACCGCATCATGSCCATATATGCGAGGTCAGCATTTTGCA 4855
DB 121 TCCCTCATCTTGCACACCGCATCATGSCCATATATGCGAGGTGCGCATTTTGCA 180
QY 4856 GTGTTCTTCTTACGGAAGTGAAGTCTCAAGGCTTTCAGTCTTCAAGTGTCTCC 4915
DB 181 GTGTTCTTCTTACGGAAGTGAAGTCTCAAGGCTTTCAGTGTCTCTCC 240
QY 4916 GCGTCTTGTGTTTACAGAGCTGCGATCTTGGTGTAACCTTGGCCGTGCGTCCG 4975
DB 241 GCGTCTTGTGTTTACAGAGCTGCGATCTTGGTGTAACCTTGGCCGTGCGTCCG 300
QY 4976 TTCTAATTGATCGGTACCGCATCATGTCGTCGCTTTGATTCGACTTTGATATTC 5035
DB 301 TTCTAATTGATCGGTACCGCATCATGTCGTCGCTTTGATTCGACTTTGATATTC 360
QY 5036 AAGGAGTTGCGTGGGCGCTGAGGTTTCTTGGGTGTTCTATGATGCTCCAGAT 5095
DB 361 AAGGAGTTGCGTGGGCGCTGAGGTTTCTTGGGTGTTCTATGATGCTCCAGAT 420
QY 5096 ATGTCATGTTCTTGGTGGGCGCTGAGGTTTCTTGGGTGTTCTATGATGCTCCAGAT 5155
DB 421 ATGTCATGTTCTTGGTGGGCGCTGAGGTTTCTTGGGTGTTCTATGATGCTCCAGAT 480
QY 5156 GCTTATGCGCTTACTTGGTGGCGCAACGCGAGATTGATCAGATCAACCGTGTCT 5215
DB 481 GCTTATGCGCTTACTTGGTGGCGCAACGCGAGATTGATCAGATCAACCGTGTCT 540
QY 5216 CCAATGCTTGCAGAAAGCAAAAGCCGAAGCAACCGCAGAAATTTTCAACGAT 5275
DB 541 CCAATGCTTGCAGAAAGCAAAAGCCGAAGCAACCGCAGAAATTTTCAACGAT 600
QY 5276 TCCACCATATCCAGGACCTTTGACCGGTGAAGCTATGSCATGAGACCGTCACGAT 5335
DB 601 TCCACCATATCCAGGACCTTTGACCGGTGAAGCTATGSCATGAGACCGTCACGAT 660
QY 5336 GCCATGTTTGCAGCGGAAAGCTTGGCTGAGTGTTCGATCGTCCCAACCAAGGGGAG 5395
DB 661 GCCATGTTTGCAGCGGAAAGCTTGGCTGAGTGTTCGATCGTCCCAACCAAGGGGAG 720
QY 5396 CTGCTTTCACAGTGAAGGAAAGATGCTGTGCTTCCCATCTGCTTCCGCA 5455
DB 721 TTAGTCTTCCCGGTGAGTGAAGATTTGTGTGCTATCCCATCTGCTTCCGCA 780
QY 5456 GTCCGACCTAAGGCTGAGATGCTTCAATGTGATATCTGTATGACATTTGGTTTCAG 5515

DB 781 GTTCCGACCAAGGCTGAGATGCTTCCATGTGATATTTGATGACATTTGGTTTCAG 840
QY 5516 ACCGTAACTTCAACGCGACGCACTTTAACCGCTGTAAGACAGGCGCATGAAGTCAAA 5575
DB 841 AAGTAACTTCAACGCGACGCACTTTAACCGCTGTAAGACAGGCGCATGAAGTCAAA 900
QY 5576 GCAAGGAGAGCTGCTGTGTAATTCGATATTGATGCTTAAAGCTTCAGATTATAGGTA 5635
DB 901 GCAAGGAGAGCTGCTGTGTAATTCGATATTGATGCTTAAAGCTTCAGATTATAGGTA 960
QY 5636 ACCAGCGGATTTGTTTGTGTAATTAAGAAACCGGACCTGTAAACCTTAAGCTTTG 5695
DB 961 ACCAGCGGATTTGTTTGTGTAATTAAGAAACCGGACCTGTAAACCTTAAGCTTTG 1020
QY 5696 GCGAATTTGAAGCGGAGCCCAACCTGTCAACGTCGCAAGAAAGACGCTGCCAGCA 5755
DB 1021 GCGAATTTGAAGCGGAGCCCAACCTGTCAACGTCGCAAGAAAGACGCTGCCAGCA 1080
QY 5756 ACACCATTAAGTTGAAAACCTTGAAGTTCG 5784
DB 1081 ACACCATTAAGTTGAAAACCTTGAAGTTCG 1109
RESULT 14
AAE71571
ID AAE71571 standard; DNA; 882 BP.
XX
AC AAE71571;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:423.
XX
XX Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
XX diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-IB000943.
XX
XX 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99US-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.

PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042086.
 PR 03-SEP-1999; 99DE-01042087.
 PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042123.
 PR 03-SEP-1999; 99DE-01042125.
 XX (BADI) BASF AG.
 XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI; 2001-061975/07.
 DR P-PSDB; AAB79454.
 XX
 PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes.
 XX
 PS Claim 3; Page 751-752; 1246pp; English.
 XX
 CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
 CC to AAB 79633 which are involved in carbon metabolism and energy
 CC production. The C. glutamicum SMP gene can be used in vectors (II) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteinogenic or
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localise C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH)
 XX
 SQ Sequence 882 BP; 259 A; 232 C; 213 G; 178 T; 0 U; 0 Other;
 XX
 Query Match 12.8%; Score 762; DB 4; Length 882;
 Best Local Similarity 93.0%; Pred. No. 1.4e-218;
 Matches 820; Conservative 0; Mismatches 60; Indels 2; Gaps 2;
 QY 1449 GTGGTCTTTGACTCAACCGCCAGGTGCAAAAGTTCATTAGTCATCACTACTTTAA 1508
 DB 1 GTGGTCTTTGACTCAACCGCCAGGTGCAAAAGTTCATTAGTCATCACTACTTTAA 60
 QY 1509 GTACGAGTAAACTATCTCTGAT-TTTAAAGAGTCCACACATGAAATCATATCTGCA 1567
 DB 61 ATACGAGCAAACTTCTCTGATTAATAAGAGTCCGACATGACATCATCTATCTGCA 120
 QY 1568 AGACGAGCAAAAGTCCGCAAAAGTTCAGTTCCTAATCGACCTTCGCAAAAGG 1627
 DB 121 AGACGAGCAAAAGTCCGCAAAAGTTCAGTTCCTAATCGACCTTCGCAAAAGG 180
 QY 1628 TGGAACTTTGGGGCTTGCAACAGATCCTACACTGAGTACTCTCAAGAGCTATTG 1687
 DB 181 CGGAACCTTTGGGGCTTGCAACAGATCCTACACTGAGTACTCTCAAGAGCTATTG 240
 QY 1688 CATGATATGAAGTGGGGAAGTGCATTCAGAAATCGCAAGGATCTTGTGGATGATA 1747
 DB 241 CATGATATGAAGTGGGGAAGTGCATTCAGAAATCGCAAGGATCTTGTGGATGATA 300
 QY 1748 CGTGGGACTTAACCCGAGCATGAAACAGCTACTTTAAACATTCGCAAGAGTTGAC 1807
 DB 301 CGTGGGATTTAAGCCGAGCATGAAACAGCTACTTTCAAAACATTCGTAAGAGTTGAC 360
 QY 1808 TGACCAATCGACATCGTTGATGAAGGCTTACAGCCCGAGTGGTGCAAACCTGATCC 1867

DB 361 TGACCAATCGACATCGTTGATGAAGAGTCTACAGCCCAATGGTGCAAAACCTGATCC 420
 QY 1868 ATACGAGCAAGTCCGAGAGTATGAGCAAAAGATCGCTCAGAAATCGTTGAAGTTCAAAT 1927
 DB 421 ATACGAGCAAGTCCGAGAGTATGAGCAAAAGATCGCTCAGAAATCGTTGAAGTTCAAAT 480
 QY 1928 CTTTGGCATCGCGGAAACGG-CACATGCTTTCACTTGAACATCATCTCTCTGTCAGG 1986
 DB 481 CTTTGGCATCGCGGAAACGG-CACATGCTTTCACTTGAACATCATCTCTCTGTCAGG 540
 QY 1987 ACTGACAAAGTCCAGGCGCTGACCCCTTAAACTGTGAGAGCAACGCTCGATCTTCAA 2046
 DB 541 ACTGACAAAGTCCAGGCGCTGACCCCTTAAACTGTGAGAGCAACGCTCGATCTTCAA 600
 QY 2047 CACCATGCAAGAGTCCCAACCCACGCGCTCAACCGGTTTGGCACTTTGTCCGCGC 2106
 DB 601 CACCATGCAAGAGTCCCAACCCACGCGCTCAACCGGTTTGGCACTTTGTCCGCGC 660
 QY 2107 GCAAAACATCGTGTGGGCAACTGTGAGAGAAAGCGACGCAATCCGCGGAACCTGT 2166
 DB 661 GCAAAACATCGTGTGGGCAACTGTGAGAGAAAGCGACGCAATCCGCGGAACCTGT 720
 QY 2167 GGAAGGCCAGTGACTGCTCTTTCGCCAGTTCATCCTGTAGATGCAACAATGCCACC 2226
 DB 721 GGAAGGCCAGTGACTGCTCTTTCGCCAGTTCATCCTGTAGATGCAACAATGCCACC 780
 QY 2227 ATCATCGTTGATGAGAGCAAGTATCCAGCTGGAACCTGTATCATCTGATCTTCAT 2286
 DB 781 CATCATCGTTGATGAGAGCAAGTATCCAGCTGGAACCTGTATCATCTGATCTTCAT 840
 QY 2287 GGAGCAATTAAGCTGGCTGGAACCAAAAGGAAGACTG 2328
 DB 841 GGAGCAATTAAGCTGGCTGGAACCAAAAGGAAGACTG 882
 XX
 RESULT 15
 AAB67871
 ID AAB67871 standard; DNA; 759 BP.
 XX
 AC AAB67871;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 2906.
 XX
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EP108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PE 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tabeiishi N, Senoh A, Ikeda M, Ozaki A;
 DR WPI; 2001-376931/40.
 DR P-PSDB; AAG92652.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 2906; 246pp + Sequence Listing; English.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 09:56:08 ; Search time 943 Seconds
(without alignments)
10357.302 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969
Sequence: 1 agtcgcgtgcagccacacat.....tcatatctgcgcgtctgaa 5969

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818139359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	61.8	6911	1	US-08-311-174-4 Sequence 4, Appli
2	2332.2	3.9	2913	4	US-09-710-279-3464 Sequence 3464, Ap
3	183.2	3.1	1452	4	US-09-489-039A-5847 Sequence 5847, Ap
4	181.4	3.0	1386	4	US-09-489-039A-6203 Sequence 6203, Ap
5	181	3.0	1491	4	US-09-134-000C-2417 Sequence 2417, Ap
6	177.6	3.0	30246	4	US-08-956-171E-56 Sequence 56, Appl
7	173.6	3.0	30246	4	US-08-781-986A-56 Sequence 56, Appl
8	173.6	2.9	1968	4	US-09-583-110-297 Sequence 297, Appl
9	172.2	2.9	3895	3	US-08-961-527-201 Sequence 201, Appl
10	167.8	2.8	1884	4	US-09-583-110-463 Sequence 463, Appl
11	167.8	2.8	1932	4	US-09-107-433-647 Sequence 647, Appl
12	162.2	2.7	1428	4	US-09-489-039A-1806 Sequence 1806, Ap
13	154.2	2.6	9769	3	US-08-961-527-30 Sequence 30, Appl
14	139	2.3	732	3	US-09-134-001C-1277 Sequence 1277, Ap
15	124.8	2.1	714	4	US-09-107-532A-378 Sequence 378, Appl
16	118.8	2.0	1905	4	US-09-543-681A-2847 Sequence 2847, Ap
17	118.6	2.0	1593	4	US-09-107-433-2566 Sequence 2566, Ap
18	117.8	2.0	951	4	US-09-134-000C-2077 Sequence 2077, Ap
19	117.6	2.0	1971	4	US-09-489-039A-1041 Sequence 1041, Ap
20	108.6	1.8	3615	1	US-08-920-812-17 Sequence 17, Appl
21	108.6	1.8	3615	1	US-08-920-827-17 Sequence 17, Appl
22	108.6	1.8	3615	1	US-08-921-177-17 Sequence 17, Appl
23	108.6	1.8	3615	1	US-08-362-577C-17 Sequence 17, Appl
24	108.6	1.8	3615	2	US-08-920-828-17 Sequence 17, Appl
25	106.4	1.8	741	4	US-09-134-000C-2317 Sequence 2317, Ap
26	104.4	1.7	804	4	US-09-489-039A-6023 Sequence 6023, Ap
27	102.6	1.7	2230	4	US-09-949-016-4239 Sequence 4239, Ap

28	99.8	1.7	1446	4	US-09-543-681A-984 Sequence 984, App
29	95	1.6	465	2	US-08-673-190A-3 Sequence 3, Appli
30	92.4	1.5	1887	4	US-09-107-532A-1634 Sequence 1634, Ap
31	92	1.5	2295	3	US-09-221-017B-736 Sequence 736, Appl
32	90.6	1.5	834	3	US-09-367-293-1 Sequence 1, Appli
33	90	1.5	1830121	4	US-09-557-884-1 Sequence 1, Appli
34	90	1.5	1830121	4	US-09-643-990A-1 Sequence 1, Appli
35	88.6	1.5	708	4	US-09-583-110-829 Sequence 829, Appl
36	87	1.5	970	4	US-09-270-767-15150 Sequence 15150, A
37	84.4	1.4	840	4	US-09-543-681A-1005 Sequence 1005, Ap
38	84.4	1.4	4403765	3	US-09-103-840A-2 Sequence 2, Appli
39	84.4	1.4	4411529	3	US-09-103-840A-1 Sequence 1876, Ap
40	73.8	1.2	1284	4	US-09-107-532A-1876 Sequence 1317, Ap
41	73.2	1.2	465	4	US-08-956-171E-1317 Sequence 1317, Ap
42	73.2	1.2	465	4	US-08-781-986A-1317 Sequence 231, Appl
43	72.2	1.2	1501	4	US-09-774-528-231 Sequence 202, Appl
44	71.2	1.2	246	4	US-09-107-433-202 Sequence 1201, Ap
45	71.2	1.2	2013	4	US-09-134-000C-1201

ALIGNMENTS

RESULT 1
Sequence 4, Application US/08311174
Patent No. 5556776
GENERAL INFORMATION:
APPLICANT: TSUCHIYA, MAKOTO
APPLICANT: MIWA, KIYOSHI
TITLE OF INVENTION: SUCRASE GENE DERIVED FROM CORNYNEFORM
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,174
FILING DATE: 23-SEP-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 046836/1992
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 5556776man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-699-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6911 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-311-174-4
Query Match 61.8%; Score 3687; DB 1; Length 6911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	3241	CTCGTGGGCTGGAATGGGGCTGCCCGCGCAGAGATGATCACCACAGTTGGACACAGAAAGA	3300
Qy	3301	TGGGTGACATGCTCCATGATCTGTGCCCGGAGCTTCATTTTGGGCAACCAAGGATCTACCA	3366
Db	3301	TGGGTGACATGCTCCATGATCTGTGCCCGGAGCTTCATTTTGGGCAACCAAGGATCTACCA	3366
Qy	3361	GAGTCTCCCTTCCCGAGAGGGGAGATCGGGGGTAAATCAGATCTGTAAATAGGTTCTGAACCT	3420
Db	3361	GAGTCTCCCTTCTCCCGAGAGGGGAGATCGGGGGTAAATCAGATCTGTAAATAGGTTCTGAACCT	3420
Qy	3421	GTCCGAGTGAACATCCGAGGCAATATTTTCCCTCGAGTGGGATGGTGTCCGTTGTCTGTG	3480
Db	3421	GTCCGAGTGAACATCCGAGGCAATATTTTCCCTCGAGTGGGATGGTGTCCGTTGTCTGTG	3480
Qy	3481	GATGCTGATGTATCTGCTCGCTAGCTGAGGTAAACCTGCGCAATAGATATGTCGGAC	3540
Db	3481	GATGCTGATGTATCTGCTCGCTAGCTGAGGTAAACCTGCGCAATAGATATGTCGGAC	3540
Qy	3541	GATTAATCAGCATTTGATGATTAACCTGAGGTATGAGCAGGTTTCATTCCGTTTCCGGGC	3600
Db	3541	GATTAATCAGCATTTGATGATTAACCTGAGGTATGAGCAGGTTTCATTCCGTTTCCGGGC	3600
Qy	3601	CTTCAAAAGTGACACTATTGAGAGATAGTATTAAGGGTCTTTTGTGTGCGCAATTTG	3660
Db	3601	CTTCAAAAGTGACACTATTGAGAGATAGTATTAAGGGTCTTTTGTGTGCGCAATTTG	3660
Qy	3661	ACAAATACCTTGCAGAAATCCCTTGATC	3687
Db	3661	ACAAATACCTTGCAGAAATCCCTTGATC	3687

RESULT 2

US-09-710-279-3464/c

; Sequence 3464, Application US/09710279

; Patent No. 6703492

; GENERAL INFORMATION:

APPLICANT: KIMBERLY, WILLIAM JOHN

TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

FILE REFERENCE: PJ3480US

CURRENT APPLICATION NUMBER: US/09/710.279

CURRENT FILING DATE: 2000-11-09

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3464

; LENGTH: 2913

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: nucleic acid sequence

US-09-710-279-3464

Qy	Query Match	3.9%;	Score 232.2;	DB 4;	Length 2913;
	Best Local Similarity	54.0%;	Pred. No. 5.8e-62;		
	Matches 574;	Conservative 0;	Mismatches 468;	Indels 21;	Gaps 4;
Db	4053	TCNAAGATGTTGTGGCTAACACGCCCAACTGTTCAACCGGCTGTGAAGTATTTGGCGG	4112		
Db	2903	TTAAAGATAAATCATCTTAAGCATATGATTCATTTCCAAAATTTGTGAAATGTTATCTG	2844		
Qy	4113	ACATTTTCGTCGCCGCTGATTCCAATCTTGTGTGTGGCGGTCTGCTCATGAGCTATACACA	4172		
Db	2843	ATACTTTGTTCCATTAATTTCCGGCTATTTGAGTGTGGTGTGTGTAATGGATTAATA	2784		
Qy	4173	ATGTGTGTGTGGCGAGATCTGTC--GTCGCGAATCACTGTGTGAGATGTTCCCTC	4229		
Db	2783	ATATTTTACACGCTAAAGATTTGTTTATGATGTAAATCATATGATGATCATATGTC	2724		
Qy	4230	AGATCAGCGGTGTCTGAGATATCAACCTGATGGCATCTGCGCCGCTTCGCTTTCG	4289		
Db	2723	AATTTTCTGGCTGTGATATGATTAATATTTTGAAGAAAGCACCTTTTACATTAATTAC	2664		

QY 4290 CAGTTTGGTGGTTTACCGCAACCAAGCTTTCCGTGCGCAATGAGTTCTGGGCGCG 4349
DB 2663 CAATACCTTATTTGGTTTATGAGCAGTAAAGCAATTTGGTGGTAACTTATTTAGGTGAG 2604
QY 4350 GCATTGTATGGGATGAGTTCCCAACCTGGTTAAAGCGCTACAGCGGTGGCGCACCA 4409
DB 2603 CTCTAGGATAGTACTTGTTCATCTCGAATGATGATGATGATGATTTTCCAAAAGCTT 2544
QY 4410 TGAACCGGCGGCA--ATGCCAATGTGCTGCTGTTGGTTGATGTTGGTCAAGCTG 4466
DB 2543 TAGAAGAAAGAAAGCTATTTCCACACTGGAGATGCTTTGGGGCTACATATTAAAGATG 2484
QY 4467 GTTACAGAGGACCGGCTTCTGTGCTGTGTCTCTTGATCTGGCAAGATCGAGA 4526
DB 2483 GTTATCAGGACAAAGATTAACCTATGCTGTGACCAATATTTAGCTACAGATTGAA 2424
QY 4527 AGTTCCTGACAAAGCACTCATGGGCACTGCAGACTTCTGATCAACCCAGTGTGACTC 4586
DB 2423 AATGCTTACGTAAAGTATTTCCACCTGTGTAGATATTTATTTAGCCCATTTATTCAA 2364
QY 4587 TCTGCTCAACGCGCTTCTTACCTTCAATGCTATTTGTCACAAATGCGCTGGGTG 4646
DB 2363 TTTTATTTACAGATTTATTAACATTTTATTTGAGGCGCTGTCACTCGTCAATTAGTT 2304
QY 4647 ACTTGCTGACACGCTGTGACAGCACTGATGATTTGGTGTGCTCAGTGGCGGTG 4706
DB 2303 ATTTGTTATCTGATGATTTGCTGTATGATTTGTTGAGCTATTTGATGATGAT 2244
QY 4707 TTTTGGCTGTGCTTACTCACCATGCTATGCTATGCTGTGACAGCAAGTCTTCCGCCA 4766
DB 2243 TATTTGTTTATTAATGCGCAATGCTATTAAGATGACACCAAGCTTTATTTAGAG 2184
QY 4767 TTGAGCTG-----GAGCTTTAACCAAGGTGATCTTCACTTTGCGAACG 4814
DB 2183 TTGAAACGACATTAATGCTGATGCACTMAAACAGTGTTCATTTATCTTCCCAATG 2124
QY 4815 CATTCATGCGCAATATCCGCGAGGGTGACAGATGTTGGCAGTGTCT---TCCTAGGA 4871
DB 2123 CAACATGCTCAATATTTGCAAAAGGTGTGACAGCTTTAGCTCAATCTTTATCATTAAGC 2064
QY 4872 AGAGTGAAGGCTCAAGGGCTTGCAGAGTGTTCAGTGTCTCCGCTGTTCTTGTGATTA 4931
DB 2063 AAAATTAATAAAGTGTGCTGCGCGGCGGATTTTCAAGTTTACAGGAATTA 2004
QY 4932 CAGACCTGCGATCTTCGTGTGAACCTTCCGCTGCGCGCTTCAATGATGATG 4991
DB 2003 CAGAACCAAGCATGTTGTGTGCAATCTTAAATGAGATATCATTTAAGGTGCTGTG 1944
QY 4992 GTACGCAAGCTATGCTGCGCTTGTGATGCACTTTGATATCAAGGAGTGTGG 5051
DB 1943 CAGGATCAGGATAGTGTGGGCTTATTTTCACTTTCAAGATTAAGGAGTGTG 1884
QY 5052 GCGCTGCAAGTCTTGGGTGTGTTCTTATGATGCTTCAGA 5094
DB 1883 GTACAGCTGATTAACCTGATTAATATATCTAATCTACACA 1841

RESULT 3
US-09-489-039A-5847
; Sequence 5847, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5847

LENGTH: 1452
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5847

Query Match 3.1%; Score 183.2; DB 4; Length 1452;
Best Local Similarity 48.4%; Fred. No. 1.2e-46;
Matches 678; Conservative 0; Mismatches 668; Indels 54; Gaps 4;

QY 3754 TTTCTCCCGGTTAAGAGAAATTCATGACCATTAAGACCTCGGCAACCATCTCG 3813
DB 57 TATATCATGATTTAGAGGATCAATGATTTTGAACAAGATTTCCGCTCACTGCTTCC 116
QY 3814 GCACATTTGGCGGGAAGCAACATTTGCGCCGCGACACTGTGCAAGCGCTTAAGCTT 3873
DB 117 CTGTGCGGCGGCAAGAAATATGCGAGCGCGCGCACTGCGCACCCCGCGGCT 176
QY 3874 CGTGTCAAAAGACCAAGATGTGATCGCCAAAGTGTGATGATGCCAGATCTGA 3933
DB 177 GGTGTGTGTCAGACGCGCTGCGGATCAGCAGGCAATGCAAAATGACGCGGTTGA 236
QY 3934 AGGCACTTTGAAACTGCGGCAATGTTCCAGATCATCTCGGCGCAGGCGATGATCA 3993
DB 227 AGGCTGTTTGCATATGCGGACAGATGATCATATCTTGGGACCGGGGTGTCAATTA 296
QY 3994 TGTTTTCAAAAGATCTGATGACGCACTCCAAAGACATGCTGTGTTCAGAGACT 4053
DB 297 AG--TCTATCCGCTTTATTCAGGCGCGAGGATCAGCAATGAGCAAAATCCGAGC 353
QY 4054 CAAGAATTTGTGCTAACAAGCAAGCACTGTGAGCGCGTGTGAAAGATTTGGCGGA 4113
DB 354 CGCGACCTGTGCGGGAAGAAAGCTGAACCGCTTCAAGGATGACCGCGCTGTCTCA 4113
QY 4114 CATTTGCTCCGCTGATTCATCTTGTGTGTGCGGTGTGCTCATGCTATCAACA 4173
DB 414 CATCTTCGTGCGATTAATTTCCGCAATGCTGTGCTCGGCTGTGATGAGGCTGTG 473
QY 4174 TGTGTGTGCGGAGATCTGTTGCTGCGGCAATCACTGTGAGATGTTCCCTCAGAT 4233
DB 474 GATGTGAAACCTTACGTTGGTGTGACCC-----GAG 506
QY 4234 CAGCGTGTGCTGATGATGATCACTGATGAGCATCTGCGCGTTCGCTTTCGCGAGT 4293
DB 507 CAACGCTCTATATCATGCTGATGTGATGTGAGATTCGCGGCGTTTATCATCTTCCGAT 566
QY 4294 GTTGTGTGTTTCAACCGCAACAAAGCTTGTGCGCATGAGTCTTGGCGCGCGGAT 4353
DB 567 CCGATCGGCTTTTACCGCGCACCGCAATTTGGCGGTAAACCTTATCTGGCGCGACCT 626
QY 4354 TGTATGCGCATGATGTTCCCAACCTGTGTTAAGGCTAAGAGTGGCGCGCACATGAC 4413
DB 627 CCGGCGGATCTTCAACCAATCCGCGCTGACCAAGCGCTGCGGCGTCTGCGC----- 677
QY 4414 CGCGGCGAAATGCAATGTGTGCTGTTGTTGATGATGTTGCTCAAGCTGTATCA 4473
DB 678 -----CGGCTTCAACACATGAATTTCTTGCGATGAGTGTGCAATGCGCTACA 731
QY 4474 GCGCACTGTCTTCTGT 4533
DB 732 GGGGACCGTCTTCCGCGT 791
QY 4534 GCACAAGGATCTCAATGAGCACTGACACTTCTGATCAACCCAGTGTGACTGTGCT 4593
DB 792 GCGCGCGTATCTCTTACGCGCTGAGACCTGATCTTCACTTCTGCTGAGGATAT 851
QY 4594 CACCGCTTCTTACGTTATGCTATGCTGTCAGCAATGCGCTGAGGTGATGCTTGT 4653
DB 852 CTCGCGCTTATTCGCGCTGTGATGCGGCGCGCGGTGCGCGCTGCGGCAAGCGCAT 911
QY 4654 GGCACAGGTGTGAGGAGATCTATGATTTGCGGTGCAAGTGTGCGGCTGTCTTTGCG 4713
DB 912 TTCGTTATCTCAGACGCTTATCAGCAAGCGCGCTGTGCGGCGCTGTGCTTTCG 971

Db 107 TGGCGATCGCCTTCGCCGTGGCCTTCG 81

RESULT 5

US-09-134-000C-2417

; Sequence 2417, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: IIS/09/134-0000

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: PatentIn version 1.0.0

SEQ ID NO 2417

LENGTH: 1491

TYPE: DNA

ORGANISM: *Er*

US-09-134-000C-2417

Query Match	3.0%;	Score 181;	DB 4;	Length 1491;
-------------	-------	------------	-------	--------------

Best Local Similarity 47.8%; Pred. No. 6.3e-46;

Matches 685; Conservative 0; Mismatches 685; Indels 63; Gaps 3

OY	4250	ATGATCAACCTGATGGCAGATCTGCGCGCGTTCGGGCTCTTGCCAGAGTGTGGTGGTTACAC	4308
Db	31	ATGATTCAGCTTAATGTCAAGCTGACACCATTTATGTTTATCCAAATCTAGTGGATTTTCC	90
OY	4310	GCAACCAAGCGTTTCGGTGGGCAAGTAGTTCTCGGCGCGCGCATTGTGTATGGCATGTG	4365
Db	91	GCTGCTAAACGTTTGTGAGCAAAACCAATTTCTAGGTGTGCAATTTGATGATGTATGACA	150
OY	4370	TTCCCAACCCGTGGTTAAACGGCTACGACGTGGCGGCACCATGACCGCGGGGCAATGCA	4422
Db	151	ACTCAGACCTTGGTGGAGCACTGA-----A	177
OY	4430	ATGATGTCCTCGTTCGTTGGATGTGCTCAAGCTGGTTACAGAGGACCGTGCCTCT	4488
Db	178	TATGGATATATTTTGGATACCATGTGTGACAAACAACTAATGCTTACCAAGTATTCGG	237
OY	4490	GTCGTGTGTGTCCTTGGAATTTGCGCAACGATCGAAGATTCCTGCAACAAGCACTATG	4544
Db	238	GTATTAGCTTCGGTATCTTTTGTCAATATTTGAAAAATATTTTCATATAAAACCTTCT	297
OY	4550	GGACATGAGACTTCCTGATCAACCCAGATGTGAATCTGCTGCTCACCGGGCTTCTTACG	4609
Db	298	TCATTTATGTGATTTTACATTCACACCATTAATATCCGTAATATTAATCTGAGATTTCTTA	357
OY	4610	TTCATTTGCTATTTGTCACGACATGCGCTGGGTGGATCTTGCTGCAACACGCTCTGACG	4665
Db	358	TTCAACGATTAATGGTCCACCAATGTTATTAATCTAATCTAAATGGAATCACAGATCTAT	417
OY	4670	GGACTCTATGATTTGGGTGGTCCAGTGGCGGCTCTGCTTTTGGGTCTGCTTACTACCA	4722
Db	418	TGGTATATTAATGCAACTGGGTCTTACATAGGAATGGGGAATCTTTGGTGGAAACAATTC	477
OY	4730	ATCGTTATCACTGCGTCTGACACGAGTCTCTCCGCAATTTAGAGCTGGAGCTGT-----	4782
Db	478	ATTGTTATGACTGCGTCTTCAATCAGTCATTTCTTGCAATCGAACAACAAGTTACTTTCAGA	537
OY	4783	-----CAACCAAGGTGATCTTTCATCTTTCGAACCGCATTCATGSCCAATATC	4831
Db	538	TGGACAAATGATGATCGACATGCGGATTTTATCTTTGATGTGCTTCATGCGCAATGTA	597
OY	4832	GCGCAGGGTGCAGCATGTTTGGCAATGTTCTTCCATAGCAAGGAAGGAAGTCAAGGCG	4899
Db	598	GCTCAGAGTGTGCACAATGTGCTATTTGGTITTTTAACTAAATAATCTTAAACAAGAGT	657

```

4992 CTTCGAGGTCCTTCAGGTGCTCCGGCTTCTTGATTAACAGAGCTCGATCTCGGT 4951
Db 658 CTGGCTTCTTCGTGTGATTAATAGACATTCCTTGATTAACAGAGCTCGATCTCGGT 717
Qy 4952 GTGAACCTTCGCTCGGCGTGGCCGTTCTCATATGATGATGATGATGATGATGATGAT 5011
Db 718 GTAAACCTTAATAACGGTTTCCCGTTTCTGTGCTCTAATGATGATGATGATGATGAT 777
Qy 5012 GCTTGTGATGACACTCTTTGATATCAAGGACGTTGCGTTGGGCGCTGACAGTTCTTGGGT 5071
Db 778 GCGATTAACAGGTCTACTTAAGTTGTTGCTGATCACTCGGTTACGCTGGAATTTCTTGGG 837
Qy 5072 GTTGTTCATATGATGTCAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 5131
Db 838 TTCCTTTCATTAACAGCACTTCTATCCCTTTTACTTCTGATGATGATGATGATGATGAT 897
Qy 5132 GTCATCGCATTCGCGCAGCGATGCTGATATGAGCTTACTTGGTTTCGCGCAGCGCAGC 5191
Db 898 TCAATTACTTTTGCAATTAATTTATTTCTATTTGATGAACAAGATCTTCCAGTATCTTGA 957
Qy 5192 ATTGATCCAGATGCAACCGCTGCTCAATGCTCTGACAGAACGACCAAGCCGACAGAA 5251
Db 958 GCAGAACCAATTCGCTGAACCAACATCAATGATGATGATGATGATGATGATGATGATGAT 1017
Qy 5252 GCACCCGCAAAATTTTCAACGATTCACACATATCCAGATATCCAGCAACCTTGAACCGGTAGCT 5311
Db 1018 GCAATATCTATGAACTTCAACACAGTAAAGAAAGATGTTGATGCTCTGCTGATGAATA 1077
Qy 5312 ATGCACTGACAGCGCTCAGCGATGCGATGTTTCCAGCGGAAAGCTTGGCTCAGGTGTT 5371
Db 1078 ACTACTCTTGGTTCAGTTATATGACCAAGATTTTCTTCAGAAATCTATAGAAAGGAGATT 1137
Qy 5372 GCGATGCTCCCAACGAGGGGCACTGTTTCAACGATGAGGGAAGATGCTGTGGTGC 5431
Db 1138 GCTATTAACCAATATGGAACACTATATATTTCTCCAGTATGATGATGATGATGATGATGAT 1197
Qy 5432 TTCCCATCTGTCACGCTTTCGACGTCGACGTCGACATGAGCTGAGGATGTTCCAATGATGAT 5491
Db 1198 TTGAAACTGAGGACGTTTATGACTCTTAATCAAAATATGAGTCT-----GAA 1245
Qy 5492 ATCTTGAATGACATATGTTTTCGACACCGTTAACTTCACAGGCGACGACCTTTAAACCGCTG 5551
Db 1246 ATATTATTTATGATGTTGTTATGACACAGTATCAATTGAATGATGAAGATTTCACTAATAA 1305
Qy 5552 AAGAAGCAGGCGCATGAAATCAAGCAGGAGCTGCTGTGTAATTCATATTTGATGCC 5611
Db 1306 GTTGTGCTCTAAACAAAAAGTAAAAAAAGGAGAGATTTCTGGAACATTTGACACGACAGTA 1365
Qy 5612 ATTAAGCTCAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5664
Db 1366 ATTAACAATTCAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1418

RESULT 6
US-08-956-171E-56
; Sequence 56, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

```



```

ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEO ID NO: 56
SEQUENCE CHARACTERISTICS:
LENGTH: 30246 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-56

Query Match      3.0%; Score 177.6; DB 4; Length 30246;
Best Local Similarity 49.1%; Pred. No. 9,5e-44;
Matches 607; Conservative 0; Mismatches 599; Indels 30; Gaps 4;

QY 3801 AACGATCTCGGCGACATTTGGCGGAGAGACAACTTGTCCGCCGACCTGTGCA 3860
DB 10819 AACCATGTAACCGTATTTGGGGAGAAAATCTTGAAGCTGACGACATTTGTAA 10878
QY 3861 CGCGTTTACGCTCGTCTCAAGAACCACAGATGTGGATGCCAAAGTCTGATGAG 3920
DB 10879 CACGATTAAGTTAGTGTGCTGAAGAGTGAAGTAAAGTTGATAAGACGATTAAGTAA 10938
QY 3921 ATCCAGATCTGAAGAGCACTTTGAAACTGGCGGACATGTTCCAGATCATGTCGGGCGAG 3980
DB 10939 ACGCTTTGGTCAAGGGGCGAGTTTAAAGCAGACATCAATCAATTCATTTGTTCCAG 10998
QY 3981 GCGATGTGGATCATGTTTCAAGAAGTCTGATGAGCAACCTCCAAAGCATTCGCTGTGT 4040
DB 10999 GACACGTGATTAAGTGTATTAAGCGATTATTATGAAACAGGTCTCAAGAGCT---T 11055
QY 4041 CCACAGAGAGCTCAAGATGTTTGGCTTAAACAAGCCAACTGTTCAAGCCGTGCTGTA 4100
DB 11056 CGAAGATGAAGCAAGAAACAGAGCTGCAAAAAGGAAATCAGTACAGATTGATGCA 11115
QY 4101 AGATTATGGCGACATTTTCCGCCGCTGATTCATCTTGGTGTGGCGGTCTGCTCA 4160
DB 11116 AATTGTGGGGATATTTTAAACCAATTAACCTGCAATGTGTGACACCTGGTTTGTAA 11175
QY 4161 TGGCTATCAACATGTGTTGGTGGCAGAGATCTGTCGGTCCGCAATCACTGGTGAGA 4220
DB 11176 TGGGATCATATATTACTTACAAATGAAGTTTATTTGGTCCAAAGCACTTATTTAGA 11235
QY 4221 TGTTCCTCAATCAAGCGGTGTGTGAGATGATCAACCTGATGAGCATTCGCGCGTTCG 4280
DB 11236 TGTATCCAAATTTGCTGATTTTCAAAATCATTAATGTGATTCGAGTACGGCATTTA 11295
QY 4281 CGTTCTCGCAGTGTGTTGGTGTTCACCGCAACCAAGCTTTCGGTGGCAATGATTTCC 4340
DB 11296 TTTTCTTACCGACATTAATGTGTGAGTAGTACGTGTATTTGGTGTAGTCCGATTC 11355
QY 4341 TGGGCGCGGCGATTGTGATGCGATGTGTTCCAAACCTGGTAAACGGCTTACAGAGTGG 4400
DB 11356 TAGGATATGTCTTAGTTGATTTTAATGCAATCCGCAATTAATCTAGATGATTTGG 11415
QY 4401 CCGCAACATGACCGCGGCGAATGCCAATGTGCTCCCTGTTTGGATGATGTTGTC 4460
DB 11416 -----CAAAAGGAATATTTCCAGCGTGAACCTTATTTGGCTTAAAGATTAAGC 11463
QY 4461 AAGCTGTATCCAGGCAACCGCTCTCTGTGCTGTGTCTCTTGTGATTTCTGGCAAGA 4520
DB 11464 AGTTGAATTTACCAAGGTCAAGTGTGCCAGTTTAAATGACGTTTACGTTCTAGATAAA 11523
QY 4521 TCGAAGATTTCTGACAAAGCACTCATGGGCACTGACAGACTTCTGATCAACCCAGTGT 4580
DB 11524 TTGAAAAGAGATTAATAAGTCGTACAGATTCATTAATAATGTTGATCGTTGAGACCG 11583
QY 4581 TGACTTCTGCTCAACGGCTTCTTACGTTCAATGCTATTGATGTCAGCAATGCGCTGGG 4640

```

```

DB 11584 TAGCCCTTTAGTACTGATTTTATGACATTTATATTCATTTGACCAAGTTCGATTATGA 11643
QY 4641 TGGGACTTCTGACACAGGTCTGACAGGACCTTATGATTTTGGTGTCAAGTCGGCG 4700
DB 11644 TTGGTACAGTATTAACATCTGTGTATCATTTATATTCACAAATGACAGGATGGCTGGCG 11703
QY 4701 GTTCCTTTTGGGTGTGTCTACACCAATGCTTATACATGCTGTGACAGACGTCCTCC 4760
DB 11704 GAGCAATATATGATTTGTTATATGACACCTTGTATTAACAGACTACCAATATGTTTT 11763
QY 4761 CGCCATTTAGCTGAGCTGTT-----CAACAGAGGTGATTCCTCATCTTCCAAACG 4814
DB 11764 TAGAGATGATTTCCAAATGATGAGTAGCACTTAAGCGGATCATTTATAGGCAATTG 11823
QY 4815 CATTCATGCGCAATATGCGGAGGCTGACAGATGTTTGGCAGTGTCTTCTTACGAAGA 4874
DB 11824 TTGGCATTTCCAAATTTGTTCAGGCTCTGACAGATTTGGAGCATGTTTGTCTATTAAC 11883
QY 4875 GT-----GAAAGCTCAAGGCGCTTGCAGGTGCTTCAGGTGCTCCGCTGTCTTG 4925
DB 11884 GTCTTAATATGTTTAAAGAGAGGCTTGGCATTAACCTTGTATTTCTGTATGTTAG 11943
QY 4926 GTATTACAGACCTTGCATCTTGTGATGAACCTTGTGCTGCGTGGCGCTTCTTACATTTG 4985
DB 11944 GTGTTACTGAACCAAGCATGTTGCTGTGAACCTTACCTGAAATATCCATTTATGCTGT 12003
QY 4986 GTATCGGTAACCGCACTATCGGTGCGCTTGTATG 5021
DB 12004 CGATATCAACCTTGTGTATTTGGGGCAATGCTTG 12039

```

RESULT 8
US-09-583-110-297
Sequence 297, Application US/09583110

```

; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: FAITH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 297
; LENGTH: 1968
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; ORIGIN: US-09-583-110-297

```

```

Query Match      2.9%; Score 173.8; DB 4; Length 1968;
Best Local Similarity 46.5%; Pred. No. 1.6e-43;
Matches 876; Conservative 0; Mismatches 962; Indels 45; Gaps 8;

QY 3797 GCGCAACGATCTGCGCAATTTGCGGAGAGCAAACTTGTCCGCGGCACTGT 3856
DB 22 GCCAAAGATCTGCTTCAAGCAATTTGAGGCAAAAGAAATGATGCTCGTAATCACTGT 81
QY 3857 GCACGCGTTTACGCTCTGCTCAAAAGACCAAGATGATGCGCCAAAGTCTGAT 3916
DB 82 GCGCAAGAGATGCGCTTGTTTTAAAGATGATTAAGAGCTATGTTAAAGCTATGAG 141
QY 3917 GATGATCGATCTGAAGAGCACTTGAACCTGGCGGATGTTCCAGATCATGTCGGG 3976
DB 142 TCAATTCAGCTGTTAAAGAACCTTTTACAAATGACAGTCAATTTTCAGTATATCAT--- 197
QY 3977 CAAAGCATGATGATCATGTTTCAAGAACTGATGACCAACTTCAAAGACATGCT 4036

```


Db 198 --TGAAATAGAGTGGCCCATCTTTTAAATGATTACAGCGCTTTCAGGATATGAGGT 255
 Qy 4037 GTGTCCAGAGAGAGTCAAAAGATTTGGTGGCTTAACAAGCCACTGGTTCAGCCCTGCT 4096
 Db 256 GTTTCAGAAAGAGCAACAGTCTGCAAGTAAGATTAACAAACGTGGTTCAAGGTGTT 315
 Qy 4097 GTGAAGATTTGGCGAGACATTTTTCGTCGGCTGATTCCAATCTGGTGGCGGTCTG 4156
 Db 316 ATGACACACTGTGGCGGAATCTTTACTCCGATTAATTCAGCCCTGATTTGTTGGTGAATG 375
 Qy 4157 CTCATGCTATCAACAATGTGTTGGTGGCGAGAT-----CTGTTGGTCCGCAATCA 4210
 Db 376 ATCTTCGGTTCCGTAATGTCTTGAAGGTGTCATTGGTCATGTTGATGGCAGAAC 435
 Qy 4211 CTGGTGAATGTTCCCTCAGATCAGCGGTGTTGCTGAATGATCACTGATGATCT 4270
 Db 436 ATCAAGAAATCCTCTAGTTTGGGAGGGGTTAACTCTCTGGTTGCTGTGTGA 495
 Qy 4271 GCGCGTTGCGCTTCTTGGCAGATGTTGGTGGTTTACCGCAACCAAGGTTTCGGTGGC 4330
 Db 496 GCTATCTTCAGTTCTTACCAAGTAGGATTACTTGGTCTGTTCTGTGAATGGGAAC 555
 Qy 4331 AATGATCTCTGGCGCCGCAATGGTATGGCGATGTTGCCAACCTGTGTTACGGC 4390
 Db 556 ACCCAAAATTTGGGAATGTTCTCGGAATCTGTTGGTATCGCTCAGTTGCTCAATGCC 615
 Qy 4391 TACAGATGCGCCGCAACCATGACCGC-----GGGGAATGCAATGTTGCTCCGTTT 4444
 Db 616 TATGGGTTGCTTCAACGCGCAGACGTGATATGCAAGCAAACTGGGTTTGAATTTGGT 675
 Qy 4445 GGTATGATGTTGCTCAAGCTGGTATACAGGGCAGCGTCTTCTGTGCTGGTGGTCTCT 4504
 Db 676 TATTTTACTGTTATATGATATGTTTACCAAGCCCAAGTATTCAGACCTTGTTCAGAGT 735
 Qy 4505 TGGATTTGGCAACGATCGAAGATTTCTGCAACAAGCACTCATAGGCACTGCAAGCTTC 4564
 Db 736 TTGAGCTGTCTTATCTGAAATTTTGGCGCAACATATCCAGAGATCAATTCATG 795
 Qy 4565 CTGATCACCCCAAGTGTGACTCTGCTGCTCACCGGCTTCTTACGTTCAATGCTATGGT 4624
 Db 796 ATTTTGTATCTTTTGTATGATGATTCAGCCCTGATTTGAGCTCATCTGCTTGGGG 855
 Qy 4625 CCAGCAATGCGTGGTGGGTGATCTGTGCAACAGCT---CTGAGGAGACTCTATGAT 4681
 Db 856 CCAATCGTGGCAATTTGACAGAGACTTTTCACTGATGTTATGGCAGGTTTAAAGGCT 915
 Qy 4682 TTCGATGCTCAGTGGCGGTCTGCTTTTCGCTGGTCTACTCAACCACTGTTTACT 4741
 Db 916 CCTGTAAATGGCTCTTGGTGAATTTTGGTGGCTTACGCTCATTTTGTATCAACA 975
 Qy 4742 GGTCTGCAACAGTCTTCCCGCAATTTGAGCTGAGCTGTT---CAACAGGCTGATTC 4798
 Db 976 GGTCTGCAACAGTCAATGCAATGCTATGATACAAATTTATGGGATCTGGTGCACAT 1035
 Qy 4799 TTTATCTTGGCAACCGCATCAATGCGCAATATGCGCAGGAGGATGTTTGGCAGG 4858
 Db 1036 GCGCTTGGCTATGATGCTCTTTTAAATATGCTCAAGGCTCAGCGCTGTTGCTAT 1095
 Qy 4859 TTTCTTCTAGCAAGAGTGAAGAGCTCAAGGCGC---TTGCAAGTGTCTTCAAGTGTCTC 4915
 Db 1096 TATTTTCAATGCGCATATGATAGTGTAGGCTCAGGTTTCACTTCTTCAACATTTTA 1155
 Qy 4916 GCTGTTCTTGTATTAAGAGCTGCGATCTTGGTGTGAACCTTGGCTGGCGGCGG 4975
 Db 1156 GCTATCTCGGTGTTACAAACAGCTCTTTTGGGGTTAAGTAAATATATTTATCA 1215
 Qy 4976 TTTCTATGATGATGCGTACGAGCTATCGAGGCGCTTGAATTTGCACTTTTATATC 5035
 Db 1216 TTTTGTGCGGATGACTGTTTCAAGCCCTTGCAGGAGCATTTATCCGTTATCTTTATGTA 1275
 Qy 5036 AAGGAGTTGCGTGGCGCTCAGGTTTCTTGGGTGTTGTTTCTATGATGCTCCAGAT 5095
 Db 1276 ACTGCGGCTTCTATGATGATCGTGGTTTGCAGGATATCTCTATATCAACTCAATAC 1335

Qy 5096 ATGTCATGTTCTTGGTTTGGCGGATTTACTTGTGATGCAATTTGGCGGAGGAT 5155
 Db 1336 ATGCTGCATTTGCAAGAACTATGCTAGTTGATGTTGTTCATGCTC-----TTG 1389
 Qy 5156 GCTTATGCGCTTACTTGGTTGGCGCAACGCGACATTTGATTCAGATCAACCGCTGCT 5215
 Db 1390 ACTTTCCTTCTTGGAGGCTGTGCTTTTCAAAATATGAGGGGATTCGAATTTGCGAG 1449
 Qy 5216 CCAAGTCTGCAAGAACGCAAGCGGAGCAAGCAACCGCAGATTTTCAACGAT 5275
 Db 1450 GCAGATTTGTTGCTCAAGAAAGACAGATTTGAGGCTATGAACAGTAGAATTTACT 1509
 Qy 5276 TCCACATCATCAAGCAACCTTTGACCGGTGAAGCTATGCACTGACAGCGCTCAGGAT 5335
 Db 1510 TCGTGAATTAATTAAGCCCACTAATCGCCAGTGAAGAAATGATCAACGACGAGT 1569
 Qy 5336 GCCATGTTTCCAGCGGAAAGCTGGCTCAGGTGTCGATGCTGCCCAAGGAGGAG 5395
 Db 1570 CCGTGTTTTGCATCAGGTGTCAATGGGGCAAGTCTGTCTATTTGAACCAAGCAGTGA 1629
 Qy 5396 CTGGTTTACCAAGTGAAGGAAAGATCGTGTGGCTTCCATCTGTGTCAGCTTTGCA 5455
 Db 1630 TTGACCTCTCCAGTTAATGAAACAGTGAAGGCTTCTTTTCCCAAGCATGCCATCGGC 1689
 Qy 5456 GTCCGCACTAAGCTGAGATGTTCCATGTCATCTTGAATGCAATTTGTTTGCAC 5515
 Db 1690 ATGTCTCTGAGAGG-----GAGTTGAATGCTCATCACTCGTATGAT 1737
 Qy 5516 ACCGTAAACCTCAAGGCAAGCACTTTAACCGCTGAAGAAACAGGAGCATGAATCA 5575
 Db 1738 AAGATGAGCTTTGATGAGCAAGTTTGAAGCTCTTGAATGCAAGAGATACGTTATA 1797
 Qy 5576 GCAGGAGAGCTGCTGTGTAATGATATGATGTCATTTAGCTGACGTTATGAGTA 5635
 Db 1798 GTTGGCAGCAACTGATTCGTTTGAATGATGATCATTAAGCTGACAGTCTGGTGA 1857
 Qy 5636 ACCAGCGGATGTTGTTTGA 5658
 Db 1858 GAACCTCTGTTATCATCA 1880

RESULT 9
 US-08-961-527-201/C
 ; Sequence 201, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,527
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 201:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3895 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-201

Query Match 2.9%; Score 172.2; DB 3; Length 3895;
 Best Local Similarity 46.5%; Pred. No. 8.9e-43;
 Matches 875; Conservative 0; Mismatches 963; Indels 45; Gaps 8;

```

QY 3797 GCGCAACGCATCTCGCGGCAATG3CGGGAAGCAACATGTCGCGCGGCACTGCT 3856
DB 3874 GCCAAGATCTGCTTCAGGCAATCGAGGCAAGAAATGTGATCGCGTAATCACTGCT 3815
QY 3857 GCAAGCGCTTACGCTCGTCTGCTCAAGACACCAAGATGCGCAAGTCTGAT 3916
DB 3814 GCGACACGAGTGGGTTGTTTGAAGATGATTAAGAGGCTTAATGTTAAAGCTATCAG 3755
QY 3917 GATGATCCAGATCTGAAAGGACCTTTGAACTGGCGGCGATGTCAGATCATCGTCGG 3976
DB 3754 TCAATTCAGCTGTTAAAGAACCTTTACAAATGACAGTCATTTCAAGTATCAT---- 3699
QY 3977 CCAGCGATGATGATCATGTTTCAAGAACTCGATGACGCACTCCAAAGACATCGCT 4036
DB 3698 --TGAAATGACGTGCCCATCTTTATATATGATTTTACAGCGGTTTCAGGATTTAGG 3641
QY 4037 GGTGTCACGACGACGCTCAAGATGTTGCTACCAACGCACTGTTGTCAGCGTCT 4096
DB 3640 GTTTCAAAGAACGACGCAAGTCTGAGTAAAGATTAACAAAGTGCTCCAGAGTGT 3581
QY 4097 GTGAAGTATGTCGCGACATTTTCGTCGCTGATTCATCTGTTGTTGTTGCGGCTCTG 4156
DB 3580 ATGACACCTCTGGCGGAATTTTAACTCGATTAATTCAGCTTATATGTCGAGGATTTG 3521
QY 4157 CTCATGCTATCAACAATGTGTTGTTGCGCAGAT-----CTGTTGCGTCCGCAATCA 4210
DB 3520 ATCCTCGTTTCCGTAATGTCTTGAAGGTGTCATGTTGTCGATTTGGATGGCAAGCC 3461
QY 4211 CTGGTGAGATGTTCCCTCAGATCAAGCGGTGTTGTCGATGATCACTGATGTCATCT 4270
DB 3460 ATCAACAATCTCTCAGTTTGGGAGGTGCAATCACTCTCTGTTGCTGTTGTA 3401
QY 4271 GCGCGCTTCGCTTCCTGTCAGTGTGTTGTTTACCGCAACCAAGCGTTTCGTTGCT 4330
DB 3400 GGTATCTTCAGTTCTTACAGATGAGGATTAATCTTGGTCTGTTCTGTAAGATGGGAAC 3341
QY 4331 AATGATCTCTGGGCGCGGCAATGGTATGCGATGTTGTTCCAACTCGTTAAAGCG 4390
DB 3340 ACCCAATTTTGGGAATGTTCTCGAATCTGTTTGTATCGCTCAGTGTCTCAATGTC 3281
QY 4391 TACGAGTGGCGCGCACCATGACCGGCGGCAAAATGCCAA-----TGTGTCCTCGTT 4444
DB 3280 TATGCGGTTGCTTCAACGCGACAGCTATATCGGCGCAACTGGGTTTGGAAATTTGCG 3221
QY 4445 GGTTCGATGTTGCTCAAGCTGTTACAGGCGACCGTCTTCCTGTCGTGTTGTTCTCT 4504
DB 3220 TATTTTACGTTAATCGATGCGTTACCAAGCCCAAGTTATCCAGCGCTTGTTCAGAGT 3161
QY 4505 TGGATTCGCGACGATCGAAGAAATTCCTGCAACAGGCACTCAATGGGCACTGCAACTTC 4564
DB 3160 TTGAGCTGCTTATCTTGAATCTTCTGCGCAAGCATATCCGAAGTCAATTTCTATG 3101
QY 4565 CTGATCAACCCAGTGTGATCTGCGTGCACCGGCTTCCTTAAGTCAATGCTATGCT 4624
DB 3100 ATTTTGTACCTTTCTGTCATGATTCAGGCTTGATTTTGGCTCATATCTTTTGGGA 3041
QY 4625 CCAGCAATGCGCTGGGTGGGTGATCTGTGCAACGCT---CTGACAGGACTTATGAT 4681

```

```

DB 3040 CCAATCGGTGAGCAATTGACAAAGACTTTCATCAGTTGTTGCGAGTTTAACGTGT 2981
QY 4682 TTCGTTGTCAGTTCGGGCTCTGTTTTCGTTGCTGCTACTACCAATGTTATCATCT 4741
DB 2980 CCAAGTAAATGCTCTTCGGGCAATTTTGGCCCTTACGCTCCATTTTGTATCA 2921
QY 4742 GATCGACACAGTCCCTCCCGCAATTTGAGGCTGTT---CAACGAGGTGATATCC 4798
DB 2920 GATTCGACCAATGACCAATGCAATGATATACAAATGATTTGGGATGCTGTGGCACT 2861
QY 4799 TTCACTTCGCAACCGCATTCATGCGCAATATCGCGAGGTTGACATGTTTGGCACTG 4858
DB 2860 GCCCATATGCCAATGATGCTCTTTCTTAATATGCTCAAGGCTGACGCTGTTGCTTAT 2801
QY 4859 TTTCTTCAGGAAGAGTGAAGAACTCAAGGCT---TTGAGAGTCTTCAAGTGTCTCC 4915
DB 2800 TATTTACATGATTCGCATGATGAGCTGAGGCTCAGTTTACTTCGCAACCATTTCA 2741
QY 4916 GCTGTTCTTGGTATTTACAGAGCTGCGATCTTCGAGTGAACCTTCGCTGCGTGGCG 4975
DB 2740 GCTTATCTCGATTTACAGAACAGCTCTTTTGGGTTAAGTAAATATATTTATCA 2681
QY 4976 TTCTAATGTTATCGTACCGACGCTATCGTGGCGCTTTGATGCACTCTTGTATTC 5035
DB 2680 TTTGTTGCTGGGATGACTGTTTCAAGCCTTCAAGGCAATGTTATCCGTTACTTTAAATG 2621
QY 5036 AAGCATGTTGCTGGGCTGCAAGTTCCTTGGTGTGTTTCTATATGATGCTCCAGAT 5095
DB 2620 ACTGGGCTTCTATTTGATGCTGTTGTCAGGATATCTCTATCAACTTCGAATAC 2561
QY 5096 ATGTCATGTTCTTGGTTTGGCGGTTAGTACTTTTTCATGATTCGAGCGGACGAT 5155
DB 2560 ATGTCGCAATTTGACAGAACTATGCTATGCTGATGTTGTTTCCATGCTC-----TTG 2507
QY 5156 GCTTATGCTTTTACTGTTTGGCGGCAACGCGACATTTATCAAGATGCAACGCTGCT 5215
DB 2506 ACTTCTCTTCCGCAAGGCTGCTCTTTTCAAAAACAGAGGCGATACCAACTTGGAG 2447
QY 5216 CCAAGTCTGACGAGAACCAAGCCGAAACAGACCGGCAATTTTCAAGAT 5275
DB 2446 GACGAATTCGTTGCTCAAGAAAGACAAATTTGAAACATGAAACGATGAACTTACT 2387
QY 5276 TCCACATCATTCAGGCACTTTTGAACGCTGAAGCTATCGACTGAGCGCTCAGCAT 5335
DB 2386 TCGGTAAATTTATCAGGCCACTAATGCGCAAGTGAAGATTTGATGATCAAGCAGAT 2327
QY 5336 GCCATGTTTGCAGCGAAAGCTTGGCTCAGGTGTTGCGATGCTCCCAAGGGGAG 5395
DB 2326 CCAATTTTTCATCAGGTGCTATGCGGCAAGGTCTATGATGAAACCAAGCCAGGTGAG 2267
QY 5396 CTGTTTCAAGAGAGGGAAGATTCGTTGGGCTTCCCATCTGATCGTCAAGCTTTGCA 5455
DB 2266 TTGACCTTCCAGTTAATGAGACAGTACGCTTCTTTCCTTACCAAGCATGCCATGCGC 2207
QY 5456 GTCCGACTAAGGCTGAGATGTTTCCATGATGATATCTTGAATGCACTTGTGTTGAC 5515
DB 2206 ATTTGCTTGAAGAG-----GATGTAATGCTCATTCACATCGGATGAT 2159
QY 5516 ACCGTAATCTCAACGCGACCACTTTAACCCGCTGAAGAAAGCAGGCGATGAAGTCAA 5575
DB 2158 ACAGTATGCTTGTATGCGCAAGGTTTGAAGTCTTGTAGTCCAAAGGATCACGTTTCA 2099
QY 5576 GCAAGGAGCTGCTGTGAAATTTGATGATTTGATGCAATTAAGGCTGAGTTATGAGTGA 5635
DB 2098 GTTGTGCAACATGATTCGTTTGTATGATGATGATTAAGGCTGCAAGGCTGTGTCACA 2039
QY 5636 ACCAGCCGATTTGTTTGGAA 5658
DB 2038 GAAACTCTGTTATCATCACCA 2016

```

RESULT 10
 US-09-583-110-463

```
/ Sequence 463, Application US/09583110
/ Patent No. 6699703
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al.
/ TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
/ TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
/ FILE REFERENCES: PAM00-07A
/ CURRENT APPLICATION NUMBER: US/09/583,110
/ CURRENT FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/107,433
/ PRIOR FILING DATE: 1998-06-30
/ PRIOR APPLICATION NUMBER: US 60/085,131
/ PRIOR FILING DATE: 1998-05-12
/ PRIOR APPLICATION NUMBER: US 60/051,553
/ PRIOR FILING DATE: 1997-07-02
/ NUMBER OF SEQ ID NOS: 5322
/ SEQ ID NO 463
/ LENGTH: 1884
/ TYPE: DNA
/ ORGANISM: Streptococcus pneumoniae
US-09-583-110-463

Query Match      2.8%; Score 167.8; DB 4; Length 1884;
Best Local Similarity 46.8%; Pred. No. 1,2e-41;
Matches 919; Conservative 0; Mismatches 962; Indels 84; Gaps 9;

QY 3779 ATGACCAATAGAGACTCTGCGGCAAGCATCTCGGCAATGCGGCGCAAGCAATTT 3838
DB 1 ATGAACAAATACAGAAATTTCAAAAAGATCATCGATCGCTTGGGCGGAGCTGAAAATGTC 60
QY 3839 GTGCGCGCGCAACACTGTGCAAGCGCTTTACGCGCTGCTCAAGACCAAGAGATG 3898
DB 61 AATAGTGTCCCACTGTGCGACTGTCTACGCTGATGATGATCAAAATGAAGAAAATC 120
QY 3899 GATCGCGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3958
DB 121 AATTAAGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 3959 TTCCAGATCATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4018
DB 181 TACCAAAATTTATCTTGGTACAGGTACAGTAAACAAATGTACGATGATGATGATGAT 240
QY 4019 ACCCTCAAGACATGCTGTGTCACAGAGAGCTCAAGAGATGTTGGCTAACAAGCGC 4078
DB 241 GCGTTACCA-----CATCATCTAAGATGATGATGATGATGATGATGATGATGATG 294
QY 4079 AACTGCTTACGCGCTGTGTAAGGATTTGGCGGCAATTTGCTCCGCTGATTTCCATC 4138
DB 295 AACTGCTTACGCGCTGTGTAAGGATTTGGCGGCAATTTGCTCCGCTGATTTCCATC 354
QY 4139 TTGCTTGTGCGGCTGTGCTCATGCTATCAACAATGTTGGTGGCGAGATCTGTTTC 4198
DB 355 ATCGTAGCGACAGGTCTCTTCAATGAGGTGCGGTGCTTTTCAACGCTTTGAAATGCCA 414
QY 4199 GGTCCGCAATCATCTGTGTGAGATGTTCCCTCAGATCAGAGGTGTTGCTGATGATCAG 4258
DB 415 CTTTCCAG-----GTGACTTTGCAACTTACACACA 444
QY 4259 CTGATGCACTGCGCGCTTGTGCGCTTGTGCGAGTGTGTTGGTTTCCAGCAACAAG 4318
DB 445 ATCTTGACAGATACAGCTTTCATCATCTTGCAGAGTTGGTTGGTGTGTCACCTTCGT 504
QY 4319 CGTTTCGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4378
DB 505 GATATTTGGGAAATCTTCGCGTGTGATGATGATGATGATGATGATGATGATGATGAT 564
QY 4379 CTGATTAAACGCTACAGAGTGGCGGCAACATGACCGGCGGGAATGCAATGATGATG 4438
DB 565 CTTCCAAACGCTTGGGCAAGTTC-----TCAAGGTGTGATGATGATGATGATGATG 615
QY 4439 CTGTTTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4498
DB 616 TCTTTTGTGTTT---CATCCTGTGTTGTTGTTGCAAGGTTCCGTTCTTCCAGCTTCA 672
```

```
QY 4499 GTCTCTTGAGATTTCTGCGCAACGATGAGAGTTCCTGCAACAAGCACTCATGCGCACTGCA 4558
DB 673 ATGCGGGGTTGTGAGAGCTAAATTTGAAAAGCTGTCCGAAAGTGTTCGATGTCATT 732
QY 4559 GACTTCTGATCAACCCAGTGTGATCTGTGCTCAACCGGCTTCTTACGTTATGCT 4618
DB 733 GACCTTGTGTAACACATGCTGTGACACTTTTGTGATGATGATGATGATGATGATGATG 792
QY 4619 ATTTGTCAGCAATGCGCTGGGTGATGATGATGATGATGATGATGATGATGATGAT 4678
DB 793 ATTTGACAGCTTTTCCAGCTGTGTAACATCACTTATGATGATGATGATGATGATGAT 852
QY 4679 GATTTTGGGTGCTGAGTGGCGGCTGCTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTG 4738
DB 853 AGCATTTCAATTTGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 912
QY 4739 ACTGCTGCAACAGTCTTCCGCAATTTGATGATGATGATGATGATGATGATGATGATG 4795
DB 913 TGAAGTGTGACCAATCTTCAATGCTTGAAGTGTGATGATGATGATGATGATGATGATG 972
QY 4796 TCCCTTCAATTTGCAACCGCATTCATGCGCAATTTGCGAGGTCAGAGATGTTTGA 4855
DB 973 AACCATTTCAAGCTATCATCAAGCTGTGATGACAGCTCAAGGTGCTGCTACTGTTGG 1032
QY 4856 GTGCTTCTTCTGCGCAAGTAAAGTCAAGGCTTTGCAAGTGTGCTGCTGCTGCTGCT 4915
DB 1033 GTTGTGTTTAAACAAATAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 1092
QY 4916 GCTGTTCTTGTATTAACAGAGCTGCTGATCTTCCGCTGATGATGATGATGATGATGAT 4975
DB 1093 GCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152
QY 4976 TTCTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5035
DB 1153 TCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
QY 5036 AAGGAGTGTGCTGCGGCGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5095
DB 1213 GCTGTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
QY 5096 ATGATCATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5155
DB 1273 CAATTTCAACATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
QY 5156 GCTTATGCTTTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5215
DB 1324 GCTTATGCTTTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
QY 5216 CAGTGTCTGCAAGAACCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5275
DB 1375 GCAAAACAAAGCTGAAAGTGTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1434
QY 5276 TCCACATCATTCAGAGCACTTTGACCGGTGAAAGTATGCACTGAGAGAGCTGAGCAT 5335
DB 1435 GAAACATTTGTA---ACTCTTATGCTCGGTGATGTTGTGCTTCTGCTGATGATGATG 1491
QY 5336 GCGATGTTTGCAGAGGAAAGCTTTGCTCAAGTGTGTCATGCTGCCACCAAGGCGAG 5395
DB 1492 CCAAGCTTCTCAAGGTGAGCTATGAGAACAGGTATGCTGTGAAACCAAGCAAGCGAG 1551
QY 5396 CTGCTTTTCAACAGTGAAGGAAAGATCGTGTGCTTCCATCTGTGTCAGCTTTTGA 5455
DB 1552 GTCTATGACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1611
QY 5456 GTCCGACTTAAGGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5515
DB 1612 TTGAAACAAAGAAATG-----GTGCTGAAGTTTGAATTCAGCTTGTGATGATG 1659
QY 5516 ACCGTTAACTTCAACGAGCAAGCACTTAAACCGCTGAAGAAAGCAAGGAGCAAGTCAA 5575
DB 1660 ACTGATCATGATGAAGGTGAGAGCTTTTGAAGCAAAAGTTGCTCAAGATGATGATGAA 1719
```

0Y	5576	GCAGGGGAGCTGCTGTGTGTAATTCGATATTGATGCCATTAAAGCTGCAAGTTATAGAGTA	5635
Db	1720	GCTGGCGCATGTTCTTGGAAACATTTTGACTCAAAACAAATCGCTGACGCTTGATGAT	1779
0Y	5636	ACCAAGCCGATGTGTTGTTTGAATTACAGAAACCGGACCTGTAAACACTTACGTTTG	5695
Db	1780	ACAACATATGTTATCGTTACAAATACAGCTACACGCTTCAGTAGCTCCAGTGGCAACA	1839
0Y	5696	GGCGAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAA	5740
Db	1840	GGTTCAGTTGCGAAGGGGATGCTGTGATCGAAGTGAATAATCTAA	1884
RESULT 11			
US-09-107-433-647			
Sequence 647, Application US/09107433			
Patent No. 6800744			
GENERAL INFORMATION:			
APPLICANT: Lynn A Doucette-Stamm and David Bush			
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID			
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC THERAPEUTICS			
NUMBER OF SEQUENCES: 5206			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: GENOME THERAPEUTICS CORPORATION			
STREET: 100 Beaver Street			
CITY: Waltham			
STATE: Massachusetts			
COUNTRY: USA			
ZIP: 02354			
COMPUTER READABLE FORM:			
MEDIUM TYPE: CD-ROM ISO9660			
COMPUTER: <Unknown>			
OPERATING SYSTEM: <Unknown>			
SOFTWARE: <Unknown>			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/107,433			
FILING DATE: 30-Jun-1998			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 60/085131			
FILING DATE: May 12, 1998			
APPLICATION NUMBER: 60/051553			
FILING DATE: July 2, 1997			
ATTORNEY/AGENT INFORMATION:			
NAME: Ariniello, Pamela Deneke			
REGISTRATION NUMBER: 40,489			
REFERENCE/DOCKET NUMBER: GTC-011			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (781)893-5007			
TELEFAX: (781)893-8277			
INFORMATION FOR SEQ ID NO: 647:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 1932 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: double			
TOPOLOGY: circular			
MOLECULE TYPE: DNA (genomic)			
HYPOTHETICAL: NO			
ANTI-SENSE: NO			
ORIGINAL SOURCE:			
ORGANISM: Streptococcus pneumoniae			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (B) LOCATION 1...1932			
SEQUENCE DESCRIPTION: SEQ ID NO: 647:			
US-09-107-433-647			
Query Match 2.8%; Score 167.8; DB 4; Length 1932;			
Best Local Similarity 46.8%; Pred. No.1,2e-41;			
Matches 919; Conservative 0; Mismatches 962; Indels 84; Gaps 9;			
3779 ATGACCATTAAGAGACTTCGCGACGACGATCTCGCGCATTTGGCGGCGAAGACAAACTT 3838			

D	49	ATGAAACAATCAGGAAATTGGCAAAAAAAGTACATGATCCTTTGGGCGGACCGTGAATAATGTCT	108
O			
O	3839	GTCCGCCCGCACACTGTGCAACGCCGTTTACGCTCGTGTCTCAAAAGCACCAAGAGATGTG	3898
D	109	AATAGTGTGGCCCACTGTGGCACTCGTCTACGTGTCAATGGTCAAAAGATGAAGAAATATC	168
O	3899	GATGCGCAAAATCTCGATGTATGATCCAGATCTGAATGAAGGACCTTTGAACCTGGCGGATG	3958
D	169	AATTAAGAAGTATGGATTGAGAACTTGGAAAAAGTTCAAGTGTCTTTTAACTCAGGGCA	228
O			
O	3959	TTCCAGATCATCGTGGGCCAGGCGATGTGATCATGTTTCAAGAATCTGATGACGA	4018
D	229	TACCAAAATTATCTTTGGTACAGTACAGTTAACAAAATGACGATGAAGTTGTGTACTT	288
O			
O	4019	ACCTTCAAAGACATCGCTGTGTCCACAGAGAGCTCAAAAGATGTGGGCTAACACGCC	4078
D	289	GGCTTACCA-----CATCATCTTAAGATGACATGAAGACGAAGTTGCTTAAACAAAGG	342
O			
O	4079	AACGTGTTCAACGCGTGTGTGAAGATATTGGCGGACATTTTCGTCGCGTATTTCCAATC	4138
D	343	AACTGGTTCCAACGTGTATCCGTATCTGTATCGGTATGTTTTCGTTCCAAATCATCTCCAGTT	402
O			
O	4139	TTGGTGTGGCGGCTTCTCTATGCGTATCAACAATGTGTGTGGTGGCAGATCTGTTC	4198
D	403	ATCTGAGGACAGAGGTCTTCTATGAGGTGGTGGTCTTTTCAACGCTCTTGAATGCCA	462
O			
O	4199	GGTCCGCAATCACTGTGTGAAGATGTTCCCTCAGATCAACGGGTGTGTGATGATGAAC	4258
D	443	CTTCCAG-----GTACCTTGGCACTTACACAA	492
O			
O	4259	CTGATGCAATCGGCGCGTTCGCGTTCTTGCGCAATGTTGGTTGTTCACCGCAACAG	4318
D	493	ATCTTGAACAGATACAGCCTTTCATCATCTTGCAGAGTTGTGTGTGTCAACTCTTCGT	552
O			
O	4319	CGTTTCGGTGGCAATGAGTGTCTGTGGCGCGGCAATTGGTATGAGATGATGTTCACAAC	4378
D	553	GTAATTTGTGAAATCTGCGCGTTGTGATCTGTTGTGTATGATGCTGTGTCTGGGTCA	612
O			
O	4379	CTGTTAACGGCTACGACGTGGCGGCAACATGACCGCGGCGAAATGCCAATGTGTCC	4438
D	613	CTTCAAAACGTTGGGCAAGTTGC-----TCAAGGTGTGAAGTAACAGCATGAAC	663
O			
O	4439	CTGTTTGGTTGAAATGTTGCTCAAGCTGGTTACAGGGCAACGAGCTTCTGTGTGGTG	4498
D	664	TTCTTTGTTT---CATCTCTGTTGTGTGTGTGCAAGTTCGTTCTTCCAGCCTTATC	720
O			
O	4499	GTCTCTTGATTTCTGGCAACGATCGAAGATTCCTGCACAAGCGACTCAGGGCACTGCA	4558
D	721	ATCGGGGTGTGCGAGCTTAAATTGAAGAAAGCTGTCCGCAAGGTGTGTTCCAGATGTCAAT	780
O			
O	4559	GACTTCTGATACACCCAGTGTGATCTGTGTCAACCGGCTTCTTAAGTCAATGCT	4618
D	781	GACCTTTGTGAACACCAATCTGTACATTTTGTGATGTATCTTATCTTGTGCACTTTGTG	840
O			
O	4619	ATTGATCAGGAATGCGCTGGGTGGGTGACTGTGTGCAACAGGCTGTGCAAGGACTAT	4678
D	841	ATTGACCAAGTTTCCAGTGTGTTGAAGAACTAATCCTTATGTGCAAAAAGCAATCTT	900
O			
O	4679	GATTTGCGTGTCCAGTCCGCGGTCTGTTTCCGTCTGTCTACTACCAATGTATC	4738
D	901	AGCATTTCAATTTGGTCTTGTGTGTGTTTCTGTATGGTGGGTTTCACAATGATGTGTGTG	960
O			
O	4739	ACTGATCGCAACAGTCTCTCCCGCAATTAGCGTGAAGCTGTTCACACAG---GTGGA	4795
D	961	TCAGGTGTGACACACATTTTCAACTGTCTTGAATGCAATTACTGTCTGTACCAATGCT	1020
O			
O	4796	TCCTTCATCTTTCGCAACCGCATCCATGGCCAAATTCGCGCAGGTGTGACATGTTGGCA	4855
D	1021	AACCATTTCAAGCTATCATCAACAGTCTGTATGACAGCTCAAGGTGTGCTACTATGTTGCG	1080
O			
O	4856	GTGTTCTTCTAAGCAAGATGAAGAAAGTCAAGGGCTTGAAGGTCTTCAAGTGTCTCC	4915
D	1081	GTGTGTGTAAACCAAAATTCCAAACTGAAGAACCTTGTCTTCCGCGCTCTCTTCT	1144

```

QY 4916 GCGTTCTTGGATTACAGAGCGCTGCGATCTTGGGTTGAACCTTGCCGCGCTGGCCG 4975
D 1141 GCGTTCTTGGATTACAGAGCGCTGCGATCTTGGGTTGAACCTTGCCGCGCTGGCCG 1200
QY 4976 TTCTACATTTGATCCGATCCGAGCTATCGGTGGCGCTTGTGATTCACCTTTGATTC 5035
D 1201 TTCTTCTTCAATGATTTGGTGCAATCGGTGGCGATTTGGCTTCTTATCTTGACATT 1260
QY 5036 AAGCAGTTTGGTGGCGCGCTGAGGTTTCTTGGGTTGTTCTATTAATGCTCCAGAT 5095
D 1261 GCTGGTACTGTAATGATGATCAACATCATCCCTGTACAAATGCTTATGTTGTTACGA 1320
QY 5096 ATGGTATGTTCTGGTGGTGGCGGATGATCCTTTGATCATTCGATGGCGGAGGATT 5155
D 1321 CAATTCACAAATCCTTCTTATGTAAGTAT-----CATTTGCCCTTGGTTT 1371
QY 5156 GCTTATGGCTTACTTCTTGGTGGCGGCAACGCGAGCATTTGATCCAGATCCAGCTGCT 5215
D 1372 GCTTCTTACTTACATGTTTGGTTACGAAG-----TGAAGTACGCAACTGCAAGCT 1422
QY 5216 CCAATGCTTGCAGAAACGAACCAAGCCGAAGACACCGCAGATTTTCAACGAT 5275
D 1423 GCAAAACAAAGCTGAAGTGGCTGAAGAAAGAAAGATTGCGCCAGACACTTTCAAAAT 1482
QY 5276 TCACATCATGATCCAGGCACTTTGACCGGTAAGCTATGCACTGACAGCGTCAAGCAT 5335
D 1483 GAAACCTTGTA--ACTCTTATCGTGGGATGTTGCTTCTTGTATGTCATTAATGAC 1539
QY 5336 GCCATGTTTGCAGCGGAAAGCTTGGCTGAGTGTGCGATCGTCCCAACCAAGGGCGAG 5395
D 1540 CCAATGTTTCTAAGTGAAGCTATGGAACAAAGATGCTGTGAACCAAGCCAGGCGG 1599
QY 5396 CTGTTTCAACAGTGAACGGAAGATGTTGGTGGCTTCCCATCTGTCAAGCTTTGCA 5455
D 1600 GTCTATGACACAGCTGATCTGAATGTTCAATTTGCTTCAACAGGCGACGCTTTGGT 1659
QY 5456 GTCCGACCTAAGGCTGAGATGTTCCATGTCGATGATCTTGAATGCAATTTGTTTGCAC 5515
D 1660 TTGAACCAAGAAATG-----GTGCTGAAGTTTGAATGATGCTGATTAATGAC 1707
QY 5516 ACCGTAACCTCAACGCGACGACCTTTAACCCGCTGAAGACGAGGCGATGATCAAA 5575
D 1708 ACTGATATCATGAACGCTGACGCTTTTGAAGCAAAAGTGTCTAAGTAAATGAGTAAA 1767
QY 5576 GCAAGGAGCTGCTGTGTAATTCATTAATGATGATGATTAAGCTGACGCTTATGAGTA 5635
D 1768 GCTGGGATGTTCTTGAACATTTGACTCAACAAATGCTGACGCTGACCTTATGATGAT 1827
QY 5636 ACCAGCCGATTTGTTTTCGATTAACAAAGAAACCGGACCTGTAAACCTTAACGTTTG 5695
D 1828 ACAACAAATGTTATCTTGAATACATGATGATGATGATGATGATGATGATGATGATGAT 1887
QY 5696 GCGCAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAAAGAA 5740
D 1888 GGTTCAGTTGCAAGGCGGATGCTGTGATGATGATGATGATGATGATGATGATGATGAT 1932

```

```

RESULT 12
US-09-489-039A-1806
; Sequence 1806, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1806

```

```

; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1806

Query Match
Best Local Similarity 47.8%; Pred. No. 5.8e-40;
Matches 612; Conservative 0; Mismatches 643; Indels 24; Gaps 4;

QY 3819 TTGGCGGCGAAGCAACATTTGGCGCGGCAACTGTGCAACGCGTTTACGCTTGTGC 3878
D 56 TGGCGGCGGCGGATGAACATCGCACCGTACAGTATTAACCGTTTGTGCTTGTGC 115
QY 3879 TCAAGACACCAAGATGATGATCCGAAAGCTGATGATGATGATGATGATGATGATGATGAT 3938
D 116 TGAACGACCGGCGCATCGCAGACCAAAAGATGAGCATGCTGCGATGTTGAAGGAT 175
QY 3939 CTTTGAACCTGCGGCGATGTTCCAGATCATGCTGCGGCGAGCGATGATGATGATGAT 3998
D 176 GCTTCACTAAGCGCGGCGATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 230
QY 3999 TCAAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4058
D 231 -CTAAGACCGCTGCTGCGGCGACACCGGCGAGACCTCCGCGCAAGAGAGGATCAAC 289
QY 4059 ATGTTGCTTCAACACCGCATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 4118
D 290 AAGCGCGCGGCAAAATATGAAGTATGAAGTATGAAGTATGAAGTATGAAGTATGAAGTAT 349
QY 4119 TCGTCCGCGTATTCATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4178
D 350 TCTTTCCTGCTGCTGCGGCGGCTGATCAGGCGGCTGATCCTCGGTTTCCGAACTGTA 409
QY 4179 TGGTTCGAGGATCTGTTGGTCCGCAATCACTGTTGATGATGATGATGATGATGATGAT 4238
D 410 TGGGAGACCTGCGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAG 469
QY 4239 GTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4298
D 470 CCATATGATCTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 529
QY 4299 TTGTTTCAACGCAACCAAGCGTTTCCGTTGATGATGATGATGATGATGATGATGATGAT 4358
D 530 TCTGCTGCTGAGCGGTTGAAGATGAGGCGGACCGCATCTTCCGGAATGCTTCTTGGG 589
QY 4359 TGGGATGTTTCCCAACCTGTTTAAAGGCTGATGATGATGATGATGATGATGATGATGAT 4418
D 590 TACCCCTGCTTCCCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649
QY 4419 GCGAAATGCAATGTTGTTTCCGTTTGGTTGATGATGATGATGATGATGATGATGATGAT 4478
D 650 -----TGTGAACTTCCGCTGTTTACATTTGCAATTTGCAAAAGTGTGCTACAGGCG 697
QY 4479 CCGTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4538
D 698 AGTATGATCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
QY 4539 AGCAGTCAATGAGCACTGAGACTTCTGATGATGATGATGATGATGATGATGATGATGAT 4598
D 818 TGTTCCTGCGACGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 877
QY 4599 GCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4658
D 818 TGTTCCTGCGACGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 877
QY 4659 ACGGTTCGAGGAGACTCT--ATGATTTGAGTGTGATGATGATGATGATGATGATGATGAT 4715
D 878 TGGGATGAGCATCTGTTTAAACGCGAGCTTCCGCGCATGCGCGCGCGCTGCTTCTGCT 937
QY 4716 TGGTCTACCAATGTTTACGTTTACGTTTACGTTTACGTTTACGTTTACGTTTACGTTTAC 4775
D 938 TCTCTACGCGCGCTGATGATTAACGCGGATGATTAACGCGGATGATTAACGCGGATGATTA 997

```

QY	4776	AGCGTTC---ACCAGGGTGGATTCCTCACTTCGACACCGCATTCATGGCCAAATTCG	4832
Db	998	AGATGATCCAGACATGAGGCGGTACCCCGGATAGGCGCTGATTCGGCTGTGCAGATTTG	1057
QY	4833	CGCAGGGTGACATGTTTGGCAGTGTCTTCTAGCGAAGATGAAAGCTCAAGGCC	4892
Db	1058	CCGAGGCTCCGGCGGTGGTGGCATCATTCATCCGACCGCGACACAGATGAAAGCTAAA	1117
QY	4893	TTGCAGGTGCTTCAGAGTGTCTCCGCTGTTCTTGAGATTAACAGACCTGGCATCTTCGGTG	4952
Db	1118	TCTGGGTACCGGCGGCAGATCTTCGGCTCATCTTCGGGGTCAACGAGCCAGCATGTACGGTA	1177
QY	4953	TGAACCTTCGCGCTGCGCTGCGCCGTTCTACAGTTGGATCGGTAACCGACACTATCGGTGGC	5012
Db	1178	TCAACCTGAAATACCGCTTCCCGATGCTGNGCGAATGATTGGGTCTGGAGCTCCGCGGCC	1237
QY	5013	CTTTCATTGCACCTCTTTGATATCAAGGACAGTGGCGTTGGCGCTGCACAGGTTCTTGGGTG	5072
Db	1238	TGCTGTGGGGCTGAACGGCGGTGCTGGGAACGGAGATGCGGTGGGGGGCTCGCGGGGA	1297
QY	5073	TTGTTTCTATTGATGCTCC	5091
Db	1298	TCTCTCCATCCAGCTTAC	1316

RESULT 13

US-08-961-527-30/c
; Sequence 30, Application US/08961527

; GENERAL INFORMATION:

APPLICANT: Charles Kunsch
TITLE OF INVENTION: Strep
INVENTOR: Charles Kunsch

```

; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:

```

STREET: 9410 Key West

;
;
;
CITY: Rockville
STATE: Maryland
COUNTY: "03

```

; COUNTRY: USA
; ZIP: 20850
; GOVERNMENT: BUREAU OF THE ARMY

```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50
; CONVERTED BY: JF-86-1007
```

```

;
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS vers
; SOFTWARE: 1.001 Beta

```

```

; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; ADDITIONAL COMMENTS: 11/06/00

```

APPLIATION NUMBER: US/08/96
FILING DATE:
CLASSIFICATION: 43A

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY / AGENT INFORMATION:

```

; AI ORNEY/AGENT INFORMATION;
;
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 35 373

```

REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB3
TELECOMMUNICATION INFORMATION:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

```

; TELEPH: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 9769 base pairs
; TYPE: nuclear acid

```

TOPOLOGY: linear
STRANDEDNESS: double
LIFE: nucleic acid

US-08-961-527-30

Query Match	2.68;	Sc
Post 10001 cfm:1 rev:1	46.78;	Sc

Best Local Similarity 46.1%; PI
Matches 917; Conservative 0;

3779 ATGACCATAGGACCTGCGC

[illegible]

Db	9710	ATGAACAATCAGAGAAATTGCAGAAAAAAGTACATGCATGCTTGGGGGAGACGTGAAGATGTC	9651
Qy	3839	GTCCGCGCCGACACTGTGCAACCGGTTTACGCTCGTGCTCAAGACACCAAGAGATGTG	3898
Db	9650	AATAGTGTGGCCACTGTGCGCACTCGTCTAAGTGTCAATGTGCTCAAGATGAAGAGAAATTC	9591
Qy	3899	GATCGCCCAATCTGAGTAGATGATCCAGATCTGMAAGGCACTTTGMAACTGGGGGCAATG	3958
Db	9590	AATAAAGAAGTGAATGAGAACTTGGAAAAAGTTCAAGGTGCTTCTTTAACTCAGGGCA	9531
Qy	3959	TTCCAGATCATCGTCGGGCGCAGCGAGTGTGATCATGTTTCAAGAACTCGATGACGCA	4018
Db	9550	TACCAAAATTACTTTGGTACAGATGACGTTAAACAAATGACGATGAAGTTGTTGTACTT	9471
Qy	4019	ACCTCCAAAGACATCGCTGTGTCCACAGACAGCTCAAAAGATGTTGTGCTPAACAGCC	4078
Db	9470	GGATTACCA-----CATCATCTTAAGGATGACATGAAGACGAAGATTTGCTAAACAGGG	9417
Qy	4079	AACGTGTTACCCGTCGTGTGAAGTAATGGCGGACATTTTCGTCGCGTATTCATC	4138
Db	9416	AACGTGTTCAACGCTATCCGTACTTTGTGTATGTTTTCGTTCAATATCCAGTT	9357
Qy	4139	TTGGTGTGACGGCTGCTGCTCATGAGTATCAACATGTTGTGTTGCGACAGATCTGTTC	4198
Db	9356	ATCGTAGGACAGAGTCTTCTATGAGTGTGGTGTGCTTTTCAACGCTTTGMAATGCA	9297
Qy	4199	GGTCGCAATCATGTGTGAGATGTTCCCTCAGATCAGCGGTGTGCTGAGATGATCAAC	4258
Db	9296	CTTCAG-----GTGACTTTGCAACTTACACAA	9267
Qy	4259	CTGATGAGATCTGGCGCGTTCGCGTCTTTCGACATGTTGTTGGTTTCAACGCAACAG	4318
Db	9266	ATCTTGAAGATAAGCCTTATCATCTTTCAGAGTTGTGTGTGTCATCACTTCGT	9207
Qy	4319	CGTTTCGGTGCATGAGTCTCTGGGCGCGGCAATGGTATGCGAGATGATGTTCCAAAC	4378
Db	9206	GTAATTTGTGAAAATTCGCGCGTTGTGTATGCTTGTGTGATGATGCTTGTCTTCGGCTCA	9147
Qy	4379	CTGTTAACGGCTACGAGCGTGGCGGCAACATGACCGGCGGCAATGCCATGTGTCC	4438
Db	9146	CTTCCAAAACGTTGGGAGTTGC-----TCAAGTGTGTGAAGTAAACGCAATGAC	9096
Qy	4439	CTGTTTGTGTAATGTTGCTCAAGCTGTGTTACAGGGCACCGTCTTCTGTGTGTGTG	4498
Db	9095	TTCTTTGTGTTT---CATCCCTGTTGTGTGTGTAAGGTTCCGTTCTCCAGCCTTATC	9039
Qy	4499	GTCCTTGGAATTCGGCAAGATGAGAAAGTTCCTGCAACAAGCACTCATGGGCACTGCA	4558
Db	9038	ATCGGGGTTTCGAGCTTAAATTGMAAAAGCTGTCGCAAGTGTTCAGATGTC-ATT	8980
Qy	4559	GACTTCTGATCACCCAGTGTGATCTGTGCTCACCCGCTTCTTACGTTGATTTGCT	4618
Db	8979	GACCTCTTGGTAACCACTTGTGACATTTTGTGTAATGTATCTTTGACCTCTTGTGC	8920
Qy	4619	ATTGCTCCAGCAATGCGTGGGTGGGTGACTGCTGGCAACGGTCTGCAAGGACTTAT	4678
Db	8919	ATTGAGACAGTTTCCAGTGTGTGAAAACCTACATCTTATTGTCAAAAAGCGATCTT	8860
Qy	4679	GATTTGCGTGTGCAAGTCGCGGCTGTGCTTTTGGGTCTGCTCATCAACAATGTTATC	4738
Db	8859	AGCATGCCATTTGCTGTGTGTGTTCTTGAATGTGTGGGTTCAACAATGATGTGTGTG	8800
Qy	4739	ACTGGTGCACAGTCTCTTCCCGCAATTGAGCTGAGGCTGTTCAACAGGG---TGA	4795
Db	8799	TCAAGTGTGACCAACATTTCAATTGCTTGAATGCAATTACTTGTGCTGACACATGCT	8740
Qy	4796	TCCTTCACTTTCGACCGCATCCATGAGCCCAATATGCGCAGGGTGAAGCATGTTGGCA	4855
Db	8739	AAACCATTCACGCTATCATCAGCTGTATGACAGCTCAAGGTGCGTCACTGTGTGCG	8680
Qy	4856	GTGTTCTTCTAGGAAGATGMAAAGCTCAAGGCGCTTTCAGAGTGTCTTCAAGTGTCTCC	4915
Db	8679	GTGTGTGTAAACAAAATAATCCAAAATGMAAACATTTGTTTCCGGGCTGCTCTTCT	8620

4916 GCTGTTCTGTATTACAGAGCTGGATCTTGGTGAACCTTGCGCTGGCGG 4975
4916 GCTGTTCTGTATTACAGAGCTGGATCTTGGTGAACCTTGCGCTGGCGG 4975
8619 GCTTCTGTATTACAGAGCTGGATCTTGGTGAACCTTGCGCTGGCGG 8560
4976 TTCTACATTTGATCGGACCGAGCTATCGGTGGCTTTGATGCACTTTGATATC 5035
8559 TTCTTCTTCAATGATGCTGGTGAATCGGTGGATGATGCTTCTTCTTGGACTT 8500
5036 AAGGCAATGCTGGGCGCTGAGTTCTTGGGTTGTTCTTATGATGCTCCAGAT 5095
8499 GCTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8440
5096 ATGTCATGTTCTGTTGGTGGGATGATGATGATGATGATGATGATGATGATGAT 5155
8439 CAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8389
5156 GCTTATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5215
8388 GCTTATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8338
5216 CAGTGTCTGCAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 5275
8337 GCAGAGAGCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8278
5276 TCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5335
8277 GAAACCTGTA---ACTCTATGCTGATGATGATGATGATGATGATGATGATGAT 8221
5336 GCATGTTTGGCAG 5395
8220 CAGTCTTCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8161
5396 CTGTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5455
8160 GTCTATGACAG 8101
5456 GTCCCACTGAAG 5515
8100 TTGAAG 8053
5516 ACCGTAACCTCAAG 5575
8052 ACTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7993
5576 GCAG 5635
7992 GCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7933
5636 ACCAG 5695
7932 ACAAAG 7873
5696 GCGCAATGAG 5740
7872 GGTTCAGTTGATGAG 7828

RESULT 14
US-09-134-001C-1277
; Sequence 1277, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1277
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (20, (22), (32))
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-134-001C-1277

Query Match 2.3%; Score 139; DB 3; Length 732;
Best Local Similarity 56.6%; Pred. No. 7.9e-33;
Matches 306; Conservative 0; Mismatches 220; Indels 15; Gaps 2;

4569 TCACCCAGTTGATCTGCTGCTACCGGCTTCTTAAGTCAATGCTATGCTCAG 4628
80 TCACCCAGTTGATCTGCTGCTACCGGCTTCTTAAGTCAATGCTATGCTCAG 139
4629 CAATGCGCTGGGTGATCTGCTGCTACCGGCTTCTTAAGTCAATGCTATGCTCAG 4688
140 TCACCTGCAATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 199
4689 GTCCAGTGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4748
200 GAGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
4749 ACCAGTCTTCCCGCAATGAGCTG-----GAGCTGTTCAACGAGGTGAT 4796
260 ACCATATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 319
4797 CCTTATCTTGGCAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4856
320 CATTTATCTTCCCATGCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 379
4857 TGTCT---TCTACGAGAGTGAAGAGCTCAAGGCGCTTGCAGTGTCT 4913
380 CATTTTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 439
4914 CCGCTGTTCTGATTAACAGAGCTGATCTTGGTGAACCTTGCTGCTGCTGCTGCTGCTGCT 4973
440 CAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
4974 GCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5033
500 CATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
5034 TCAAGGAGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5093
560 TAAAGCATGAGCGCTTGTACAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGA 619
5094 A 5094
620 A 620

RESULT 15
US-09-107-532A-378
; Sequence 378, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 378:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 714 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (8) LOCATION 1..714
 SEQUENCE DESCRIPTION: SEQ ID NO: 378:
 US-09-107-532A-378

Query Match 2.1%; Score 124.8; DB 4; Length 714;
 Best Local Similarity 52.5%; Pred. No. 2.5e-28;
 Matches 380; Conservative 0; Mismatches 322; Indels 22; Gaps 4;

QY 1544 CACCATGGAATCTACTATCTGCAAGAGCAGCAAGAGTGG--CAAGCAGTTGCACT 1600
 DB 9 CACTATGGAATCATTCGTGTAAGAAAACCTGAGAGGCGGAAAAAGCATTGAAAT 68
 QY 1601 CCTATCGACCCCTCGCCAAAGGCTGCACTTGGGGCTTGCACAAGATCCTCAAC 1660
 DB 69 CATCAAGAGGATGAGCAAGAGTGCAGAAAGTATTAAGATTGGCAACAGAGCACACC 128
 QY 1661 ACTGAGTACTACCAAGAGCTCATTCGCATGTATGAGAGCTGGGAAAGTGTCAATCAAGAA 1720
 DB 129 AGTTACTTATATAAGAAATGACTTCA-----GCGATTGGATTTCAGTGA 176
 QY 1721 CTGCAAGGCAATCTGTGGATGAAATACGTGGACTAACCCGTGACGATGAACAAGCTA 1780
 DB 177 AATGACTTCTGTAACCTAGATGAGTACGTGACTAGCGGTTCAGATGAACAAGCTA 236
 QY 1781 CTTTAAACCATTCGCAAGAGTTCACCTGACATCGACATCGATGAGAGAGTCTA 1840
 DB 237 CCGTCACTTTCATGATGATGTCATTAATTTGATATAAAAAACCATTTCA-----AAGAAACGTT 290
 QY 1841 CAGCCAGATGAGTGAACCCGTATCCATACAGAGAGCTGCAAGATATAGAGCAAGAT 1900
 DB 291 TGTACCAAGTGAAGAAAGCAAGAGATCTGAGACAGATCAAAAAGTATGATCAATCAT 350
 QY 1901 CGCTGCAAGATCCGTTGAAGTTCAAAATCTTGGCATCGCGGAAACGG--CACATGGCTTT 1959
 DB 351 CGATGCAATCCAAATCGATATCCAAATCTTAGAAATCGGGCAAAATGACATATCGGTTT 410
 QY 1960 CATTGAACCATCATCTCTCTGTCAGAGTGAACAAGATCAAGGCGCTGCACCTATAAAC 2019
 DB 411 CAATGAACCTGGTGTCTTCAATTGACGAAAGCACTTCAGTAGTTGACTTAACGAATCAAC 470
 QY 2020 TGTGAGGACAAAGCTCGATTTCTTCAACACATCGAAGAGTCCCAACCAAGCGGTCAAC 2079

DB 471 AATCAATCAAAACAAAGCATTTCTTGATTAAGCAGAGATGTACCAACAAAGCTGTTTC 530
 QY 2080 CCAGGTTTGGGCACTTTGTCCCGCGCAAAACATGCTGTTGGTGGCACTGGTGAAG 2139
 DB 531 AATGGAAATCGGTTCAATTATGAAGAAATGCTGTTGATGGCATTTGGCGAAAG 590
 QY 2140 AAAAGCGAGCCATCCGCGGAGACTGTGAGAGGCGGAGTGAAGTCTGCTTCTGCCAGGTTTC 2199
 DB 591 CAAGCTGATCAATCAAGGAGATGTAAGACCTGTAACGAATCATTTGCCAGCAAG 650
 QY 2200 CATCTGTAGTGCACAACATGACCAATCATCTGTTGATGAAGACAGCATATCAAGCT 2259
 DB 651 TGTATTGCAAAATCACGACATGTGTTGTGATCATGATGAAGAGCAGCAAGTAAGTT 710
 QY 2260 GGAA 2263
 DB 711 GTAA 714

Search completed: March 9, 2005, 13:51:46
 Job time : 960 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 12:11:23 ; Search time 2980 Seconds
(without alignments)
11885.615 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969

Sequence: 1 agtccgcgtacgcaccacat.....tgatatttcgcgcctgaa 5969

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5401638 seqs, 2966923429 residues

Total number of hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	5656.2	94.8	3309400	9	US-09-738-626-1
2	1874.2	31.4	1963	9	US-09-738-626-2904
3	1412.6	23.7	1527	17	US-10-450-055-9
4	1321.2	22.1	1342	18	US-10-781-014-347
5	1267.8	21.2	1299	9	US-09-738-626-2905
6	1237.4	20.7	1287	18	US-10-781-014-413
7	1237.4	20.7	1287	18	US-10-781-014-415
8	1110.4	18.6	1152	9	US-09-738-626-2907
9	762	12.8	882	18	US-10-781-014-423
10	671.8	11.3	759	9	US-09-738-626-2906
11	294	4.9	1543	19	US-10-494-541-9

ALIGNMENTS

RESULT 1

US-09-738-626-1/c

Sequence 1, Application US/09738626

Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIRO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NAKO

APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO

APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738, 626

CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 1

LENGTH: 3309400

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

US-09-738-626-1

Query Match 94.8%; Score 5656.2; DB 9; Length 3309400;

Best Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3

QY	1	AGTCGGTGAAGCCCAACATTAATGTGTGTGTACCCAGACTCCGGAGGCTTTTCAATCT	60
Db	2815585	AGTCGGTGAAGCCCAACATTAATGTGTGTGTACCAAGAACTTGGAGGCTTTTCAATCC	2815528
QY	61	ACGCTCCCGTCGGCCGTGAGTGGGGTCATTACGGGTGGATCACGCCGTGAAGTTGCG	120
Db	2815525	ACACTGCCGTGGCGGTGAGTGGGGTCATTACGGGTGGGATCACGCCGGTGAAGTTGCG	2815466
QY	121	GAACCCATGTGTCTCTTGTGGGTGAGGAAACGATGCCGGTGAAGACTTTTCAAGTC	180
Db	2815465	GAAGCCATGTGTCTCTTGTGGGTGAGGAAACGATGCCGGTGAAGACTTTTCAAGTC	2815406
QY	181	TCTGCAGTTTAAATTAATGATCATCATCAGCTTGGAAGCCTAGGTAATTCAGTAGACTG	240
Db	2815405	TCTGCAGTTTAAATTAATGATCATCATCAGCTTGGAAGCCTAGGTAATTCAGTAGACTG	2815346
QY	241	CAACGACAGGCTCAAGTCCGAGATTAATTAAGTATCCGTGACATTAAGACATCATTA	300
Db	2815345	CAACGACAGGCTCAAGTCCGAGATTAATTAAGTATCCGTGACATTAAGACATCATTA	2815286
QY	301	CGTCCATATGCTTGCTGGAAGAACCAATTAACCTAGAAATATGGCAGATGGGCACTT	360
Db	2815285	CGTCCATATGCTTGCTGGAAGAACCAATTAACCTAGAAATATGGCAGATGGGCACTT	2815226
QY	361	ATCAAGAAATATGCAAGTCAAGCATTTAAAAAATTTGAGGAAGATTTGTTCCCCCTCG	420
Db	2815225	ATCAAGAAATATGCAAGTCAAGCATTTAAAAAATTTGAGGAAGATTTGTTCCCCCTCG	2815166
QY	421	GGGTATTTGATGGCTTTCTCCACATCGAAAAACGGATATATGACGGAATCTCTGAGAAC	480
Db	2815165	GGGTATTTGATGGCTTTCTCCACATCGAAAAACGGATATATGACGGAATCTCTGAGAAC	2815106
QY	481	CAGCACCTAAAAACGACAGATTTCAACCCCGAATCTCCACGATTTGTTCCCGTTTATYG	540
Db	2815105	CAGCACCTAAAAACGACAGATTTCAACCCCGAATCTCCACGATTTGTTCCCGTTTATYG	2815046
QY	541	ATCTTCATATATCAAGGTGGGAAACGGTGGGGGTTTCCATAGGGAAACGACGACACGCGA	600
Db	2815045	ATCTTCATATATCAAGGTGGGAAACGGTGGGGGTTTCCATAGGGAAACGACGACACGCGA	2814986
QY	601	GGAACACCGCGCATATCACCCGCGAACTGGCAACGACCCGTATGTTGCAAGCATGGTTT	660
Db	2814985	GGAAATGCGCGCCACATATCACCCGCGAACTGGCAACGACCCGTATGTTGCAAGCATGGTTT	2814926
QY	661	CGGGCGGGGCTGACGCACTGGCAGGGCAGGTGGAAAACTTAATTCCTGTGTGTAACAAG	720
Db	2814925	CGGGCGGGGCTGACGCACTGGCAGGGCAGGTGGAAAACTTAATTCCTGTGTGTAACAAG	2814866
QY	721	TCTGTGTGTGCGGCAATTCACCTCGAAGGACCCTTTCATCAACGCAATGCGTGTGTGCTC	780
Db	2814865	GCTGTGTGTGCGGCAATTCACCTCGAAGGACCCTTTCATCAACGCAATGCGTGTGTGCTC	2814806
QY	781	AAAAACCCGGATTTCAATTTTCCCGGCAACCAACAGATCTTGCCCGGTATCCATGCGG	840
Db	2814805	AAAAACCCGGATTTCAATTTTCCCGGCAACCAACAGATCTTGCCCGGTATCCATGCGG	2814746
QY	841	GAAGAAAGTTGATCAAAATCATATCAAGTAGGCGCGGAAACCTGAACATCTTCTAGCTTC	900
Db	2814745	GAAGAAAGTTGATCAAAATCATATCAAGTAGGCGCGGAAACCTGAACATCTTCTAGCTTC	2814686
QY	901	TTCGATCTCTGGCAGCGCACCATCATATGCTTCTCTTGCGGCAACATGATGACATTTTG	960
Db	2814685	TTCGATCTCTGGCAGCGCACCATCATATGCTTCTCTTGCGGCAACATGATGACATTTTG	2814626
QY	961	ATTACCACTACAGCGCAATTCCTTGGCTAAAGAAAAATGTACCGGTACCGCTACGCG	1020
Db	2814625	ATTACCACTACAGCGCAATTCCTTGGCTAAAGAAAAATGTACCGGTACCGCTACGCG	2814566
QY	1021	ATTGTGTCAATGCAATGCTTCGCTGCATCATAGGAGCTCCCGACGCTGCGGCTTTTGC	1080

Dd	2814565	ATTGTTCAATGCGATGCTTCCTCGTCATCTAGGAAATCCCGGCAAGCGTGCGCCCTTTGC	2814560
Qy	1081	TTGCTGCGGCAAGTCCCGGGACGCAATATGTTGAATTGATCGCCGACGCGTGCATTTGG	1140
Dd	2814505	TTGCTGCGGCAAGTCCCGGGACGCAATATGTTGAATTGATCGCCGACGCGTGCATTTGG	2814466
Qy	1141	CCGATGGAACGCTGATCTGATGCTGTTCCAAACAAGCCTTTTATCATACGAGAGCCCATGG	1200
Dd	2814445	CCGATGGAACGCTGATCTGATGCTGTTCCAAACAAGCCTTTTATCATACGAGAGCCCATGG	2814388
Qy	1201	AAGCCGCGGAATCCAGACGCGTAGTACATATTTTGCGCGTTTGAACGTACACCTGCACCG	1260
Dd	2814385	AAGCCGCGGAATCCAGACGCGTAGTACATATTTTGCGCGTTTGAACGTACACCTGCACCG	2814326
Qy	1261	ATGGAGTTCGCCCTGCTCTGCGGATATGGCGGCGCCATTCGCCGGGGGACCAAGACATATACGA	1320
Dd	2814325	ATGGAGTTCGCCCTGCTCTGCGGATATGGCGGCGCCATTCGCCGGGGGACCAAGACATATACGA	2814266
Qy	1321	GTCAGTTTCGATCCACAGTGGCGAGGGGATATGACGCTTATTCAGAGCGACCTCCACACT	1380
Dd	2814265	GTCAGTTTCGATCCACAGTGGCGAGGGGATATGACGCTTATTCAGAGCGACCTCCACACT	2814206
Qy	1381	CAACCGTCGCGCTAAATTTCTCGTCTTGCGATCCGAATTCGCTAAATCCAACTCG	1440
Dd	2814205	CAACCGTCGCGCGCAAAATTTCTCGGACTTACGAAATTCGTTAAATCCAACTCG	2814146
Qy	1441	CAAAATTTTGAGTCTTTGACCTCAACGCGGACGAGGCAAAAGTGCATTTAGTGCATCAG	1500
Dd	2814145	TAATTTTGTGAGTCTTTGACTCAACGCGGACGCTTACCAAGGCTCATTTTGAACCATCAG	2814086
Qy	1501	TACTTTAAGTACGAGTAAATCTATCTGAT-TTTAAAGAGTCCACCATGGAATCACT	1559
Dd	2814085	TAATTTAATATACGAGCAAACTTTCTGTATATATTAAGAGTCCGACCATGAGCATATC	2814026
Qy	1560	ATCTGCAAAACGAGCAAGAGTCCGAAAGCGATGSCATCTCTAAATCCGACCTTGCC	1619
Dd	2814025	ATCTGCAAAAGCGAGCAAGAGTCCGAAAGCGGACGCGACCTGATCGACCTTTGCA	2813966
Qy	1620	AACAAGGATGGAACTTTGAGGCTTGCAACAGAGATCTTCAACATGAGTACTTACCAAG	1679
Dd	2813965	ACTAAGGCGGAACTTTGGGGCTTGCAACTGGAATCGTCACTTTGAGACCTTACCAAG	2813906
Qy	1680	CTCATTTGCGATGTATGAAGCTGGGGAAAGTGTCAATTCAAGAACTGCAAGGCAATTCTTGG	1739
Dd	2813905	CTCATTTGCGATGTATGAAGCTGGGGAAAGTGTCAATTCAAGAACTGCAAGGCAATTCTTGG	2813846
Qy	1740	GATGAATACGTGGGACTTAACCCGTCAGATGAAGAAACAGCTACTTTAAACCATTCGCAAA	1799
Dd	2813845	GATGAATACGTGGGACTTAACCGCGAGCATGAAGAAACAGCTACTTTCAAAACCATTCGTA	2813786
Qy	1800	GAGTTCACTGACCATCGACATCTGTTGATGAAGAGGTCTACAGCCCAATGTTGTCAAC	1859
Dd	2813785	GAGTTCACTGACCATCGACATCTGTTGATGAAGAGGTCTACAGCCCAATGTTGTCAAC	2813726
Qy	1860	CCGATATCATACGAAGAGCGTGCAGAGTATGAGGCAAAAGATCGTGCAGAAATCGTTGAA	1919
Dd	2813725	CTGATATCATACGAAGAGCTGTCAGAGTATGAGGCAAAAGATCGTGCAGAAATCGTTGAT	2813666
Qy	1920	GTTTCAAAATCTTTGGCATCGCGGAAACG-CACATCGCTTTCAATTGAACCATCATCTTCT	1978
Dd	2813665	GTTTCAAAATCTTTGGCATCGCGGAAACGCGCACATCGCTTTCAATGAGCATCATCTTCT	2813606
Qy	1979	CTGTCAAGAGTGCAGAAAGTTCAGGGCGTGCACCTTAAACTGTGAGAGCAACGCTCGA	2038
Dd	2813605	CTGTCAAGAGTGCAGAAAGTTCAGGGCGTGCACCTTAAACTGTGAGAGCAACGCTCGA	2813546
Qy	2039	TTCTTCAACACCAATCGAAGAGTCCCAACCAACGCGTCAACCAAGGTTTGAGCACTTTG	2098
Dd	2813545	TTCTTCAACACCAATCGAAGAGTCCCAACCAACGCGCTTCAACCAAGGTTTGAGCACTTTG	2813486
Qy	2099	TCCGCGCGCAAAACATCGTGTGTTGTGGCAACTGTGTGAAGAAAGCCGACGCATTCGCG	2158
Dd	2813485	TCCGCGCGCAAAACATCGTGTGTTGTGGCAACTGTGTGAAGAAAGCCGACGCATTCGCG	2813426

QY 2159 GGAACCTGAGGAGCCAGTGAAGTCTTCTGCTCCAGGTTCCATCTCTGATGACACAC 2218
 |||||
 Db 2813425 GGAACGTGAGGAGCCAGTGAAGTCTTCTGCTCCAGGTTCCATCTCTGATGACACAC 2813366
 QY 2219 ATGCCACATCATCTGATGAGGAGGATGATCCAGCTGAGAAACGCTGATCACTAC 2278
 |||||
 Db 2813365 ATGCCACATCATCTGATGAGGAGGATGATCCAGCTGAGAAACGCTGATCACTAC 2813306
 QY 2279 GGTCTCATGAGCAATTAAGTCTGCTGAGAAACAAAGAAAGTACTGTGTGGGCTA 2338
 |||||
 Db 2813305 GGTCTCATGAGCAATTAAGTCTGCTGAGAAACAAAGAAAGTACTGTGTGGGCTA 2813246
 QY 2339 TGCACACAAATCTTCCAGTTTGGCCCTGCGTACATGATGATCTCTCCGACGGGACAGC 2398
 |||||
 Db 2813245 TGCACACAAATCTTCCAGTTTGGCCCTGCGTACATGATGATCTCTCCGACGGGACAGC 2813186
 QY 2399 TCAATGATCCCAACGGAATGATAGTATGATGAGATATACCTCCACGCTACTACACACAG 2458
 |||||
 Db 2813185 TCAATGATCCCAACGGAATGATAGTATGATGAGATATACCTCCACGCTACTACACACAG 2813126
 QY 2459 ATCCAGGTTCCCTTCGACCAACAGCGACCGGCTGGCTCACACACACACCGCGTTGA 2518
 |||||
 Db 2813125 ATCCAGGTTCCCTTCGACCAACAGCGACCGGCTGGCTCACACACACACCGCGTTGA 2813066
 QY 2519 CCGGACCGGACGATGTCAGTGAAGCGACCTGCGGACGCTCTTTACCCGATGATCCT 2578
 |||||
 Db 2813065 CCGGACCGGACGATGTCAGTGAAGCGACCTGCGGACGCTCTTTACCCGATGATCCT 2813006
 QY 2579 ATGACCTGATGATGATCTATTCGCGTGAAGCGGATTTTACTGACGACACTTAACCTT 2638
 |||||
 Db 2813005 ATGACCTGATGATGATCTATTCGCGTGAAGCGGATTTTACTGACGACACTTAACCTT 2812946
 QY 2639 TCTACACCGGCAACCTTAATTTGACGAAAGCGCGCCACCCAAACCTTGTGAAG 2698
 |||||
 Db 2812945 TCTACACCGGCAACCTTAATTTGACGAAAGCGCGCCACCCAAACCTTGTGAAG 2812886
 QY 2699 TCGAGACCCCAACTGGGCTGATGGGCGCATTCATGCGGCTTCCGCTTAACCTTCCGCTTA 2758
 |||||
 Db 2812885 TCGAGACCCCAACTGGGCTGATGGGCGCATTCATGCGGCTTCCGCTTAACCTTCCGCTTA 2812826
 QY 2759 TCGAGACCCCGCAGCGGTTTCAACCCCATTAACCGGATCCCATGATCAGGCTGATG 2818
 |||||
 Db 2812825 TCGAGACCCCGCAGCGGTTTCAACCCCATTAACCGGATCCCATGATCAGGCTGATG 2812766
 QY 2819 GTGATGTTGGAACATGTTCTTGGGCGCCAAACGCGAAACCTCACCGGTGACGCGTTG 2878
 |||||
 Db 2812765 GTGATGTTGGAACATGTTCTTGGGCGCCAAACGCGAAACCTCACCGGTGACGCGTTG 2812706
 QY 2879 TATACCGCTCGACATCTTGAACCTGGAATTTCCGCTGAAATCACCCTTGACCTCA 2938
 |||||
 Db 2812705 TATACCGCTCGACATCTTGAACCTGGAATTTCCGCTGAAATCACCCTTGACCTCA 2812646
 QY 2939 GTGATGACCAACCTGTTCTGCTGCTGATCTGTTCCGATGGCTCATGTGGGAATGTC 2998
 |||||
 Db 2812645 GTGATGACCAACCTGTTCTGCTGCTGATCTGTTCCGATGGCTCATGTGGGAATGTC 2812586
 QY 2999 CCAACCTTTTACGCTTCCGATGAGAAACCTGCGAAGATCTCGACGTTGATTTTCT 3058
 |||||
 Db 2812585 CCAACCTTTTACGCTTCCGATGAGAAACCTGCGAAGATCTCGACGTTGATTTTCT 2812526
 QY 3059 GTTCACAAAGATTTGACCCGATTCACAGATGAGTTACTCATACGCAAGCTCTGACCAAGT 3118
 |||||
 Db 2812525 GTTCACAAAGATTTGACCCGATTCACAGATGAGTTACTCATACGCAAGCTCTGACCAAGT 2812466
 QY 3119 GCGGATGATGCTGCAAGCTTGAAGAAACGACCTTCCGCTGCTGCGAGGATTCAGCG 3178
 |||||
 Db 2812465 GCGGATGATGCTGCAAGCTTGAAGAAACGACCTTCCGCTGCTGCGAGGATTCAGCG 2812406
 QY 3179 AGCTGATTTCCGCAATGATTTCTACGACCCGAGGTTGCAAGTAAACGCTTCTGATGCT 3238
 |||||
 Db 2812405 AGCTGATTTCCGCAATGATTTCTACGACCCGAGGTTGCAAGTAAACGCTTCTGATGCT 2812346

QY 3239 GGTGCTGGGGCTGATGGGGCTGGCCGCGACAGATGATATACCAACAGTTGACAGAGAG 3298
 |||||
 Db 2812345 GGTGCTGGGGCTGATGGGGCTGGCCGCGACAGATGATATACCAACAGTTGACAGAGAG 2812286
 QY 3299 GATGGGTGACCTGCTGATCTGTGCCCCGACGCTTCAATTTGGCAACACAGGATCTAC 3358
 |||||
 Db 2812285 GATGGGTGACCTGCTGATCTGTGCCCCGACGCTTCAATTTGGCAACACAGGATCTAC 2812226
 QY 3359 AAGAGCTCTTCTCCACAGGGGGAGTGGGGGATATCATGATCTGATTTAGTTCTGAA 3418
 |||||
 Db 2812225 AAGAGCTCTTCTCCACAGGGGGAGTGGGGGATATCATGATCTGATTTAGTTCTGAA 2812166
 QY 3419 CTGTCCGATGACATCCGAGGCAATTTCCCTGAGGGAGTGGTGTGCGTTGCTG 3478
 |||||
 Db 2812165 CTGTCCGATGACATCCGAGGCAATTTCCCTGAGGGAGTGGTGTGCGTTGCTG 2812106
 QY 3479 TGGATCTGATGATGATCTGCTGCGTGAAGTAAACCTGCGCAATTAAGTATGATCGCG 3538
 |||||
 Db 2812105 TGGATCTGATGATGATCTGCTGCGTGAAGTAAACCTGCGCAATTAAGTATGATCGCG 2812046
 QY 3539 ACGATATACAGCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3598
 |||||
 Db 2812045 ACGATATACAGCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2811987
 QY 3599 GCTTCAAGGATGACATTAATGAGATGATGATGATGATGATGATGATGATGATGATGAT 3658
 |||||
 Db 2811986 GCTTCAAGGATGACATTAATGAGATGATGATGATGATGATGATGATGATGATGATGAT 2811927
 QY 3659 GTACAAATCTCTGCAAAATCCCTTATGTCGACACAAATTAACAGTTTAATTTGTTTA 3718
 |||||
 Db 2811926 GTACAAATCTCTGCAAAATCCCTTATGTCGACACAAATTAACAGTTTAATTTGTTTA 2811867
 QY 3719 GCTTTTGAACAAATCTCATGTCGATATTTTGTCTTCTCCGCTTAAGGAAATTC 3778
 |||||
 Db 2811866 GCTTTTGAACAAATCTCATGTCGATATTTTGTCTTCTCCGCTTAAGGAAATTC 2811807
 QY 3779 ATGACCAATTAAGGATCTGCGGACGATCCGCGGACATTTGGGGCGGACGACATTT 3838
 |||||
 Db 2811806 ATGACCAATTAAGGATCTGCGGACGATCCGCGGACATTTGGGGCGGACGACATTT 2811747
 QY 3839 GTGCGCGCGGACATGTCGCAACGCGTTTACGCTGTCGTCGCAAGACCAAGATGTC 3898
 |||||
 Db 2811746 GTGCGCGCGGACATGTCGCAACGCGTTTACGCTGTCGTCGCAAGATGTC 2811687
 QY 3899 GATCGCAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3958
 |||||
 Db 2811686 GATCGCAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2811627
 QY 3959 TTCCAGATCATGTCGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 4018
 |||||
 Db 2811626 TTCCAGATCATGTCGCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 2811567
 QY 4019 ACCCTCAAAAGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 4078
 |||||
 Db 2811566 ACCCTCAAAAGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2811507
 QY 4079 AACTGATGACCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 4138
 |||||
 Db 2811506 AACTGATGACCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2811447
 QY 4139 TTGATGTCGCGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 4198
 |||||
 Db 2811446 TTGATGTCGCGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2811387
 QY 4199 GGTGCGCAATCACTGATGAGATGTTTCCCTCAGATCAGCGGTTGCTGAGATGATCAAC 4258
 |||||
 Db 2811386 GGTGCGCAATCACTGATGAGATGTTTCCCTCAGATCAGCGGTTGCTGAGATGATCAAC 2811327
 QY 4259 CTGATGCAATCTGCGCGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4318
 |||||
 Db 2811326 CTGATGCAATCTGCGCGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2811267
 QY 4319 GCTTTCCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4378

Db 2811266 CTTTTGGGGGCAATGATGTTCTGGGGCGCGGATTTGGATGATGCGATGTTTCCGAGC 2811207
Qy 4379 CTGGTTAACGCGTACGACGTCGCGCCGACCATGACCGCGGCGAAATGTCATGTCGTC 4438
Db 2811206 TTGGTGAACGGCTACGACCTGGCCGACCATGCTGCGGGCGAAATGTCATGTCGTC 2811147
Qy 4439 CTGGTTGGTTGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4498
Db 2811146 CTGGTTGGTTGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2811087
Qy 4499 GTCTCTGGATTTCTGGCAACGATGCAAGATTTCTGCAACAGCACTGAGGCACTGCA 4558
Db 2811086 GTTCTTGGAATTTCTGGCAACGATGCAAGATTTCTGCAACAGCACTGAGGCACTGCA 2811027
Qy 4559 GACTTCTGATACCCCAAGTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4618
Db 2811026 GACTTCTGATACCCCAAGTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2810967
Qy 4619 ATTGGTCCAGCAATGCGTGGGTGGGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4678
Db 2810966 ATTGGTCCAGCAATGCGTGGGTGGGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2810907
Qy 4679 GATTTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4738
Db 2810906 GATTTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2810847
Qy 4739 ACTGCTGCTGCAACGATGCTTCCCGCAATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 4798
Db 2810846 ACTGCTGCTGCAACGATGCTTCCCGCAATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 2810787
Qy 4799 TTTCATCTTGCAGCAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4858
Db 2810786 TTTCATCTTGCAGCAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2810727
Qy 4859 TTCTTCTAGCGAAGATGAAAAAGCTCAAGGCGCTTGCAGGCTGCTGCTGCTGCTGCTGCT 4918
Db 2810726 TTCTTCTAGCGAAGATGAAAAAGCTCAAGGCGCTTGCAGGCTGCTGCTGCTGCTGCTGCT 2810667
Qy 4919 GTTCTTGGTATTAACAGAGCTGCGATCTTGGTGTGAACTTCGCTGCGCTGCGCTGCTGCT 4978
Db 2810666 GTTCTTGGTATTAACAGAGCTGCGATCTTGGTGTGAACTTCGCTGCGCTGCGCTGCTGCT 2810607
Qy 4979 TACATTTGATGCGTACCGGACGCTATCGGTGGCGCTTATTTGCACTCTTTGATATCAAG 5038
Db 2810606 TTCAICGCGATGCGTACCGGACGCTATCGGTGGCGCTTATTTGCACTCTTTAATATCAAG 2810547
Qy 5039 GCAAGTTGCGTGGCGCTGCAAGGTTCTTGGGTGTTGTTCTATTGATGCTCCAGATATG 5098
Db 2810546 GCAAGTTGCGTGGCGCTGCAAGGTTCTTGGGTGTTGTTCTATTGATGCTCCAGATATG 2810487
Qy 5099 GTCACTGCTTGGTGGTGGCGGCTGATGCTTGTGATGCGATGCGGCGACCGATGCT 5158
Db 2810486 GTCACTGCTTGGTGGTGGCGGCTGATGCTTGTGATGCGATGCGGCGACCGATGCT 2810427
Qy 5159 TATGCGCTTACTTGGTGGCGCAACGCGACATGATCCAGATGCAACCGCTCTCCA 5218
Db 2810426 TATGCGCTTACTTGGTGGCGCAACGCGACATGATCCAGATGCAACCGCTCTCCA 2810367
Qy 5219 GTGCGCTGAGAACGACCAAGCCGAGAGAGACCCGCGAGATTTTCAAGCATTC 5278
Db 2810366 GTGCGCTGAGAACGACCAAGCCGAGAGAGACCCGCGAGATTTTCAAGCATTC 2810307
Qy 5279 ACCATCATCCAGACGACCTTTGACCGGTGAAGCTATGCGCTGAGAGGCTGACGCGATGCG 5338
Db 2810306 ACCATCATCCAGACGACCTTTGACCGGTGAAGCTATGCGCTGAGAGGCTGACGCGATGCG 2810247
Qy 5339 ATGTTTGCAGCGAGAAAGCTTGGCTCAGGCTGTTGGATGCTCCCAACCAAGGGGAGCTG 5398
Db 2810246 ATGTTTGCAGCGAGAAAGCTTGGCTCAGGCTGTTGGATGCTCCCAACCAAGGGGAGCTG 2810187
Qy 5399 GTTTCACGATGAGCGGAAAGATGCTGTGCGCTTCCCATCTGATCAACGCTTTCGAGTC 5458

Db 2810186 GTTCTCCGCGTGAAGTGAAGATGTTGGGCAATTCATCTGCGCATGCTTTCGAGTT 2810127
Qy 5459 CGCATTAAGGCTGAGAGATGTTTCCATGTCATCTTGAATGACATGTTTTCGACACC 5518
Db 2810126 CGCACCAAGGCTGAGAGATGTTTCCATGTCATCTTGAATGACATGTTTTCGACACA 2810067
Qy 5519 GTPAACCTCAACGCGACGACCTTTAACCGCTGAGAGAGAGGCGATGAGTCAACACA 5578
Db 2810066 GTPAACCTCAACGCGACGACCTTTAACCGCTGAGAGAGAGGCGATGAGTCAACACA 2810007
Qy 5579 GGGAGCTGCTGTGTAATTCGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 5638
Db 2810066 GGGAGCTGCTGTGTAATTCGATTTGATGATGATGATGATGATGATGATGATGATGAT 2809947
Qy 5639 ACGCGATGTTGTTTTCGAAATTAAGAAACCGGACCTGTAAACCTTACGTTTGGGC 5698
Db 2809946 ACGCGATGTTGTTTTCGAAATTAAGAAACCGGACCTGTAAACCTTACGTTTGGGC 2809887
Qy 5699 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTCAGCAACA 5758
Db 2809886 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTCAGCAACA 2809827
Qy 5759 CCATTAAGTTGAAACCTTGAAGTGTGCAACACAGTTAGCTAGGGAGCTGACCTCTACGC 5818
Db 2809826 CCATTAAGTTGAAACCTTGAAGTGTGCAACACAGTTAGCTAGGGAGCTGACCTCTACGC 2809767
Qy 5819 ATCTTGAACACCGGATACCGGATACGTTGATGATTTTAAACCTGTTCAACAGGTCATGCC 5878
Db 2809766 ATCTTGAACACCGGATACCGGATACGTTGATGATTTTAAACCTGTTCAACAGGTCATGCC 2809707
Qy 5879 TCGGTGATCCTGTGTGTGTCACACCCGCAATCTTCAACCCCATTTGACATGTTGTTCA 5938
Db 2809706 TCGGTGATCCTGTGTGTGTCACACCCGCAATCTTCAACCCCATTTGACATGTTGTTCA 2809647
Qy 5939 GCAAGTACGTTGATATTTTGGCGCGCTG 5967
Db 2809646 GCAAGTACGTTGATATTTTGGCGCGCTG 2809618

RESULT 2
US-09-738-626-2904
Sequence 2904, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIRO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US-09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 2904
LENGTH: 1983
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-2904

Query Match 31.4%; Score 1874.2; DB 9; Length 1983;

Best Local Similarity 96.6%; Pred. No. 0;
Matches 1915; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

QY 3779 ATGGACCATPAAGAGACCTCGCGCAACGATCTCGCGACATTTGGCGCGCAAGCAACATT 3838
DB 1 ATGGACCATPAAGAGACCTCGCGCAACGATCTCGCGACATTTGGCGCGCAAGCAACATT 60
QY 3839 GTGCGCGCGCACACTGTGCAACGCGTTTACGCGCTGTCTCAAAAGCACACAGATGTG 3898
DB 61 GTGCGCGCGCACACTGTGCAACGCGTTTACGCGCTGTCTCAAAAGCACACAGATGTG 120
QY 3899 GATCGCCAAAGTCTGATGATGATTCAGATCTGAAAAGCACCTTTGAAACTGCGCGCATG 3958
DB 121 GATCGCCAAAGTCTGATGATGATTCAGATCTGAAAAGCACCTTTGAAACTGCGCGCATG 180
QY 3959 TTCCAGATCATGCTCGCGCGCGCGAGATGATGATCATGTTTCAAAAGACCTTCATACGCA 4018
DB 181 TTCCAGATCATGCTCGCGCGCGCGAGATGATGATCATGTTTCAAAAGACCTTCATACGCA 240
QY 4019 ACCTCAAAAGACATCGCTGTGTCCACAGAGACACTCAAAAGATGTGTGCTAAACGCGC 4078
DB 241 ACCTCAAAAGACATCGCTGTGTGTCCACAGAGACACTCAAAAGATGTGTGCTAAACGCGC 300
QY 4079 AACTGTTAGCGGTGCTGTGAAGGTAATGCGGACATTTTCTCCGCTGATTCATC 4138
DB 301 AACTGTTAGCGGTGCTGTGAAGGTAATGCGGACATTTTCTCCGCTGATTCATC 360
QY 4139 TTGGTTGGTGGCGGTCTGTCTCATAGGCTATCAACATGTGTGTTGCGCAGATCTGTTC 4198
DB 361 TTGGTTGGTGGCGGTCTGTCTCATAGGCTATCAACATGTGTGTTGCGCAGATCTGTTC 420
QY 4199 GGTCCCAATCACTGTGTGAGATGTTCCCTCAGATCAGCGGTGTGTCTGATGATCAAC 4258
DB 421 GGTCCCAATCACTGTGTGAGATGTTCCCTCAGATCAGCGGTGTGTCTGATGATCAAC 480
QY 4259 CTGATGGAATCTGCGCGCTTCCGCTTCTTGCAAGTGTGTGTTTCAACCGCAACCAAG 4318
DB 481 CTGATGGAATCTGCGCGCTTCCGCTTCTTGCAAGTGTGTGTTTCAACCGCAACCAAG 540
QY 4319 CGTTTCGGTGGCAATAGTTCCTCGGCGCGCGGATGTTGATGCGCATGTTTCCCAAC 4378
DB 541 CGTTTCGGTGGCAATAGTTCCTCGGCGCGCGGATGTTGATGCGCATGTTTCCCAAC 600
QY 4379 CTGGTTAAGCGCTACGACGTGCGCGCACCATGACCGCGGCGAAATGCAATGTGTTC 4438
DB 601 TTGGTGAAGCGCTACGACGTGCGCGCACCATGACCGCGGCGAAATGCAATGTGTTC 660
QY 4439 CTGTTTGGTTGGATGTGCTCAAGCTGTGTTACAGAGGACCGTGTCTTCTGTGCTGTG 4498
DB 661 CTGTTTGGTTGGATGTGCTCAAGCTGTGTTACAGAGGACCGTGTCTTCTGTGCTGTG 720
QY 4499 GTCTCTGATGATCTGCGCAACGATCGAAGATCTCTGCAACAGGACTCATGCGCACTGCA 4558
DB 721 GTCTCTGATGATCTGCGCAACGATCGAAGATCTCTGCAACAGGACTCATGCGCACTGCA 780
QY 4559 GACTTCCTGATACCCGAGTGTGACTGTGCTGCTACCGGCTCTTACGTTCAATTTGCT 4618
DB 781 GACTTCCTGATACCCGAGTGTGACTGTGCTGCTACCGGCTCTTACGTTCAATTTGCT 840
QY 4619 ATTGTGCAAGATGCGCTGTGAGTGAATTGCTGCGCAACGATCTGCAAGGACTCTAT 4678
DB 841 ATTGTGCAAGATGCGCTGTGAGTGAATTGCTGCGCAACGATCTGCAAGGACTCTAT 900
QY 4679 GATTTGCGTGTGCAAGTGTGCGGTCTGCTTTTGTGTCTACTACCAATCGTTATC 4738
DB 901 GATTTGCGTGTGCAAGTGTGCGGTCTGCTTTTGTGTCTACTACCAATCGTTATC 960
QY 4739 ACTGCTGTGCAAGATCTTCCGCGCAATGAGTGTGAGTGTGCAACGAGGTGTGATCC 4798
DB 961 ACTGCTGTGCAAGATCTTCCGCGCAATGAGTGTGAGTGTGCAACGAGGTGTGATCC 1020
QY 4799 TTCTATCTTGGCAACCGCATCAATGCGCAATATGCGCGAGGTGCAAGCATGTTTGGCAGTG 4858

```

```

DB 1021 TTCTATCTTGGCAACCGCATCTATGCTAATATGCGCCAGGATGCGCATGTTTGGCAGTG 1080
QY 4859 TTCTTCTTGGCGAAGGATGATAAGCTCAAGGCGCTTGGAGGCTTCAAGTGTCTCCGCT 4918
DB 1081 TTCTTCTTGGCGAAGGATGATAAGCTCAAGGCGCTTGGAGGCTTCAAGTGTCTCCGCT 1140
QY 4919 GTTCTTGTATTTACAGAGCTCGATCTTCTGAGTGAACCTTTCGCTGCGTGGCGGCTTC 4978
DB 1141 GTTCTTGTATTTACAGAGCTCGATCTTCTGAGTGAACCTTTCGCTGCGTGGCGGCTTC 1200
QY 4979 TACATTTGATCGGTAACGCGACCTATCGGTGCGCTTGTGATGCACTTTGATATCAAG 5038
DB 1201 TACATTTGATCGGTAACGCGACCTATCGGTGCGCTTGTGATGCACTTTGATATCAAG 1260
QY 5039 GCAATGTGCTTGGCGCTGCAAGTGTCTTGGGTGTGTTTCTATGATGCTCAAGTATG 5098
DB 1261 GCAATGTGCTTGGCGCTGCAAGTGTGTTGTTGTTCTATGATGCTCAAGTATG 1320
QY 5099 GTCATGTTCTTGGTTTGGCGGTAAGTTACCTTTGTATGCAATTCGCGCAGCATTTGCT 5158
DB 1321 GTCATGTTCTTGGTTTGGCGGTAAGTTACCTTTGTATGCAATTCGCGCAGCATTTGCT 1380
QY 5159 TATGCTTTTACTTGTGTGCGGCAACGCGACATTTGATCAAGTGAACCGCTGTCCA 5218
DB 1381 TATGCTTTTACTTGTGTGCGGCAACGCGACATTTGATCAAGTGAACCGCTGTCCA 1440
QY 5219 GTGCTGCAAGAACGACCAACCGAAGCAAGCAACCGCGAATTTTCAACGATTC 5278
DB 1441 GTGCTGCAAGAACGACCAACCGAAGCAAGCAACCGCGAATTTTCAACGATTC 1500
QY 5279 ACCATCATCAGGACCTTTGACCGGTGAAGTATGCACTGAGACGCTGACGATGCG 5338
DB 1501 ACCATCATCAGGACCTTTGACCGGTGAAGTATGCACTGAGACGCTGACGATGCG 1560
QY 5339 ATGTTTGCAGCGAAGAGCTTGTGCTCAAGTGTGCGATGCTCCCAACAAAGGCGAGCTG 5398
DB 1561 ATGTTTGCAGCGAAGAGCTTGTGCTCAAGTGTGCGATGCTCCCAACAAAGGCGAGCTG 1620
QY 5399 GTTTTACCAAGTGAAGGAAAGTGTGTGCTGCTTCCATCTGTGTCAACCTTTCCGAGTC 5458
DB 1621 GTTTTACCAAGTGAAGGAAAGTGTGTGCTGCTTCCATCTGTGTCAACCTTTCCGAGTC 1680
QY 5459 CGCATTAAAGGCTGAAGATGTTCCATGATGATCTTGAATGATGCAATGTTTCCGACCC 5518
DB 1681 CGCATTAAAGGCTGAAGATGTTCCATGATGATCTTGAATGATGCAATGTTTCCGACCC 1740
QY 5519 GTTAACTTCAACGCGACCTTTAACCGCTGAAGAAAGCAAGGCGAGTGAAGTCAAGCA 5578
DB 1741 GTTAACTTCAACGCGACCTTTAACCGCTGAAGAAAGCAAGGCGAGTGAAGTCAAGCA 1800
QY 5579 GGGAGCTGCTGTGTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5638
DB 1801 GGGAGCTGCTGTGTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 5639 ACGCGATTTGTTTGTGTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 5698
DB 1861 ACGCGATTTGTTTGTGTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1920
QY 5699 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTGCAAGCA 5758
DB 1921 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTGCAAGCA 1980
QY 5759 CCA 5761
DB 1981 CCA 1983

```

RESULT 3
US-10-450-055-9
; Sequence 9, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft

```

; TITLE OF INVENTION: No. US20040043953A1el genes of Corynebacterium
; FILE REFERENCE: 936, 2000
; CURRENT APPLICATION NUMBER: US/10/450,055
; CURRENT FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 9
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1504)
; OTHER INFORMATION: KX500315
; US-10-450-055-9

Query Match      23.7%; Score 1412.6; DB 17; Length 1527;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

QY 4259 CTGATGCACTGCGCGCTTCCGCTTCTTGCCAGTGTGTTGTTTCAACGCAACGAAG 4318
DB 1 CTCATGCACTGCGCGCTTCCGCTTCTTGCCAGTGTGTTTCAACGCAACGAAG 60
QY 4319 CGTTTCGGTGGCAATGATTCCTGGGCGCG--GCATTGGTATGGCGATGTGCCAAC 4377
DB 61 CGTTTCGGGCGCAATGATTCCTGGGCGCGGCTATGGATGGCATGTGTCCGAG 120
QY 4378 CCTGTAAACGGCTACGACGTGGCCGCCACATGACCGCGGCGGAATGCCAATGTGTC 4437
DB 121 CTGTGGAAAGGCTACGACGTGGCGGCCACATGCGTGGCGGGAATGCCAATGTGTC 180
QY 4438 CCTGTGTTGGTATGATGTTCTCACTGCTGTTTCAAGGCAACCGTCTCTGTGCTGT 4497
DB 181 CCGTGTGGTATGATGTTGCCCAAGCCGCTTACAGGCGCACCGTCTCTGTGCTGT 240
QY 4498 GGTCTCTTGATTCGGCAACGATGGAAGTTCGCAACGCGCATCATGGGCACTGCG 4557
DB 241 GGTCTCTTGATTCGGCAACGATGGAAGTTCGCAACGCGCATCATGGGCACTGCG 300
QY 4558 AGACTTCTGATCAACCCGATGTTGACTGCTGCTCACCGCTTCTTACGTTACCTTGC 4617
DB 301 AGACTTCTGATCACTCCAGTGTGCTGCTGCTCACCGATTCCTTACATTCATGCG 360
QY 4618 TATGTGTCAGCAATCGCTGGTGGTGTACTTCTGCGCACAGCTGTGCAAGGACTCTA 4677
DB 361 CATGGCCAGCAATCGCTGGTGGTGTACTTCTGCGCACAGGCTGTCAAGGACTCTA 420
QY 4678 TGATTTGGGATGCGATGCGGCGGTGCTGCTTTCGCTGATGCTACTGCAACAATGCTAT 4737
DB 421 TGATTTGGGATGCGATGCGGCGGTGCTGCTTTCGCTGATGCTACTGCAACAATGCTAT 480
QY 4738 CACTGCTGCAACCAAGTCTTCCCGCAATTGAGCTGAGCTGTTCAACCAAGGTGATC 4797
DB 481 CACTGCTGCAACCAAGTCTTCCCGCAATTGAGCTGAGCTGTTCAACCAAGGTGATC 540
QY 4798 CTTCATCTTCCGCAACCGCATTCATGCGCAATATCGCGAGGGTGGCAAGTGTGGCACT 4857
DB 541 CTTCATCTTCCGCAACCGCATTCATGCGCAATATCGCGAGGGTGGCAAGTGTGGCACT 600
QY 4858 GTTCTTCCAGCAAGAGTAAAGCTCAAGGGCTTGGCAGGTGCTTCAAGGTGTCTCCCG 4917
DB 601 GTTCTTCCAGCAAGAGTAAAGCTCAAGGGCTTGGCAGGTGCTTCAAGGTGTCTCCCG 660
QY 4918 TGTTCTTGATTAACAGAGCTGCGATCTTCCGCTGTAACCTTCCGCTGCGCTGCGCTT 4977
DB 661 TGTTCTTGATTAACAGAGCTGCGATCTTCCGCTGTAACCTTCCGCTGCGCTGCGCTT 720
QY 4978 CTTCATCTTGGTATGCGCAACCGCATTCGCTGGCGCTTGGATGCACTCTTATATCA 5037
DB 721 CTTCATCTTGGTATGCGCAACCGCATTCGCTGGCGCTTGGATGCACTCTTATATCA 780
QY 5038 GGCAGTGGCGTTGGGCGCTGCAAGTTCCTTGGGTGTTGTTTATGATGCTCCAGATAT 5097

```

```

DB 781 GGCAGTGGCGTTGGGCGCTGCAAGTTCCTTGGGTGTTTCTATGATGCTCCAGATAT 840
QY 5098 GGTCAATGTTCTTGGTTTGGCGGATAGTACCTTGTGTCATTCGATTCGGCGACGATTCG 5157
DB 841 GGTCAATGTTCTTGGTGTGCAAGTGTACCTTCTTCACTGCAATTCGGCGACGATTCG 900
QY 5158 TTATGCGCTTTTACTTGTGTTCCCGCAAGCGCATTTGATCCAGATGCAACCGCTGCTCC 5217
DB 901 TTATGCGCTTTTACTTGTGTTCCCGCAAGCGCATTTGATCCAGATGCAACCGCTGCTCC 960
QY 5218 AGTGCCTGAGAAAGCAACCAAGCCGAAGCAGAAAGCAACCCGACGATTTCAAGCATTC 5277
DB 961 AGTGCCTGAGAAAGCAACCAAGCCGAAGCAGAAAGCAACCCGACGATTTCAAGCATTC 1020
QY 5278 CACCATCATCCAGGACCTTTGACCGGTGAGACTATGCACTGAGCAGCGTACGATGC 5337
DB 1021 CACCATCATCCAGGACCTTTGACCGGTGAGACTATGCACTGAGCAGCGTACGATGC 1080
QY 5338 CATGTTTCCAGCGGAAGCTTGGCTCAGTGTGGCATGCTGCCACCAAGGGGCACT 5397
DB 1081 CATGTTTCCAGCGGAAGCTTGGCTCAGTGTGGCATGCTGCCACCAAGGGGCACT 1140
QY 5398 GGTTCACAGTGAACGGAAGATGCTGAGCTTCCATCTGCTCAACGCTTTCGCACT 5457
DB 1141 AGTTTCTCCGATGATGGAAGATTTGTGTGGCATTCATCTGCGCATGCTTTCGCACT 1200
QY 5458 CCGCATTAAGGCTGAGATGTTCCATGATGATATCTTATGCACTTGGTTTCGACAC 5517
DB 1201 TCGCACCAAGGCTGAGATGTTCCATGATGATATCTTATGCACTTGGTTTCGACAC 1260
QY 5518 CGTAACCTCAACGGCAAGCACTTAAACCGCTGAAGAGAGAGGCGATGAAAGTCAAGC 5577
DB 1261 AGTAACCTCAACGGCAAGCACTTAAACCGCTGAAGAGAGAGGCGATGAAAGTCAAGC 1320
QY 5578 AGGGAGCTGCTGTGTAATGATATGATGCACTTAAGCTGCAAGTATGAGATAC 5637
DB 1321 AGGGAGCTGCTGTGTAATGATATGATGCACTTAAGCTGCAAGTATGAGATAC 1380
QY 5638 CAGCGCGATGTTGTTTGAATTAACAAGAAACCGGACCTGTAAACACTTACGTTTGGG 5697
DB 1381 CAGCGCGATGTTGTTTGAATTAACAAGAAACCGGACCTGTAAACACTTACGTTTGGG 1440
QY 5698 CGAAATGGAAGGGAGCAACCTGCTCAAGTGCAGAAAGAAAGAGCGGTGCCAGAAC 5757
DB 1441 CGAAATGGAAGGGAGCAACCTGCTCAAGTGCAGAAAGAAAGAGCGGTGCCAGAAC 1500
QY 5758 ACCATAGTTGAACCTTGAAGTTCG 5784
DB 1501 ACCATAGTTGAACCTTGAAGTTCG 1527

RESULT 4
US-10-781-014-347
; Sequence 347, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; INVOLVED IN CARBON METABOLISM AND ENERGY
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: EGI-126CPN
; CURRENT APPLICATION NUMBER: US/10/781,014
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 09/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208

```


PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 601/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remainder Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 784
SEQ ID NO 347
LENGTH: 1342
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1342)
OTHER INFORMATION: RXN00041
US-10-781-014-347

Query Match 22.1%; Score 1321.2; DB 18; Length 1342;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2238 ATGAAGCAGCATGATTCAGAGCTGGGAAAGCGTGATACCTGCTCATGAGGCAATTAA 2297
DB 1 ATGAAGCAGCATGATTCAGAGCTGGGAAAGCGTGATACCTGCTCATGAGGCAATTAA 60
QY 2298 AGCTGCGCTAGAAACAAAGAAAGTAAGTGTGGGGCTATGCAACAGAACTTTCCAG 2357
DB 61 AGCTGCGCTAGAAACAAAGAAAGTAAGTGTGGGGCTATGCAACAGAACTTTCCAG 120
QY 2358 TTTCGGCCCTGCGTACATGTGAATCTTCGCGAGGGGAGGCTCAATGATCCCAAGGAAT 2417
DB 121 TTTCGGCCCTGCGTACATGTGAATCTTCGCGAGGGGAGGCTCAATGATCCCAAGGAAT 180
QY 2418 GATAGCTGATGAGATACCTCTCAAGCTTACTTACGAGCATCCAGGTTTCCCTTCCG 2477
DB 181 GATAGCTGATGAGATACCTCTCAAGCTTACTTACGAGCATCCAGGTTTCCCTTCCG 240
QY 2478 ACCAAGCGCAGCGGCTGGGCTCAGACACAGCGGTTGACCGGACCGGAGGATTTGA 2537
DB 241 ACCAAGCGCAGCGGCTGGGCTCAGACACAGCGGTTGACCGGACCGGAGGATTTGA 300
QY 2538 GTGAGCGCAGCTGCGCGAGCTCTTTACCGGATGATGATCTGATGATGATGATGCTA 2597
DB 301 GTGAGCGCAGCTGCGCGAGCTCTTTACCGGATGATGATCTGATGATGATGATGCTA 360
QY 2598 TTTCGGTGAAGCCGTATTTACTGACGCGCACCTTAACTTTTCTAACCGGCAACTTAA 2657
DB 361 TTTCGGTGAAGCCGTATTTACTGACGCGCACCTTAACTTTTCTAACCGGCAACTTAA 420
QY 2658 AATTGACGAAAGCGCGCGCCACCAAAACCTTGTGGAAGTGAAGACCACTGGGCT 2717
DB 421 AATTGACGAAAGCGCGCGCCACCAAAACCTTGTGGAAGTGAAGACCACTGGGCT 480
QY 2718 GATGGCGGCAATTCATGCGCTTGTGCTTAAATTCGCTTATGACGAGCCGCGCGG 2777
DB 481 GATGGCGGCAATTCATGCGCTTGTGCTTAAATTCGCTTATGACGAGCCGCGCGG 540
QY 2778 TTTCACACCCCATTTACCGGATCCATGATCAGCCCTGATGATGATGATGATGATGAT 2837
DB 541 TTTCACACCCCATTTACCGGATCCATGATCAGCCCTGATGATGATGATGATGATGAT 600
QY 2838 TCTGGGGGCGCAACGCGAAACCTGACCGGTTGAGGCTTCTATACGCTCGACATGCT 2897

DB 601 TCTTGGGGCCCAACGCGAAACCTCACCGGTGACGCGTGTATACCGCTCGACAGATCT 660
QY 2898 TGAACACCTGGGAATTCCTCGGTGAAATACCTTTGACCTCGATGATGACAACTGGTTC 2957
DB 661 TGAACACCTGGGAATTCCTCGGTGAAATACCTTTGACCTCGATGATGACAACTGGTTC 720
QY 2958 TGTCTCTGATCTGCTTCCGAGTGGCTATCATGATGGAATGCCCCAACCTTTTACGCTTCG 3017
DB 721 TGTCTCTGATCTGCTTCCGAGTGGCTATCATGATGGAATGCCCCAACCTTTTACGCTTCG 780
QY 3018 CGATGAAGAAACCTGCGAAGATCTTCGACGCTGTATTTTCTGTCACAGATTTGACCG 3077
DB 781 CGATGAAGAAACCTGCGAAGATCTTCGACGCTGTATTTTCTGTCACAGATTTGACCG 840
QY 3078 AATCAGATGAGTGTACTCATGACGATCTTGAACAGGCGGATATGTCGTGACAA 3137
DB 841 AATCAGATGAGTGTACTCATGACGATCTTGAACAGGCGGATATGTCGTGACAA 900
QY 3138 GCTTGAAGAAACGACCTTCCGCTCTTGCAGAGATTCAGCAGCTGATTTTCGCGCATGA 3197
DB 901 GCTTGAAGAAACGACCTTCCGCTCTTGCAGAGATTCAGCAGCTGATTTTCGCGCATGA 960
QY 3198 ATTCTACGACCGCAGGTTGAGTAACGATTCTGATGCTGCTGCTGCTGCTGATGAG 3257
DB 961 ATTCTACGACCGCAGGTTGAGTAACGATTCTGATGCTGCTGCTGCTGCTGATGAG 1020
QY 3258 GCTGCGCGCGCAGATGATGATCAACCAAGTTGCAAGAGATGAGTGGTCACTGGCTGAC 3317
DB 1021 GCTGCGCGCGCAGATGATGATCAACCAAGTTGCAAGAGATGAGTGGTCACTGGCTGAC 1080
QY 3318 TGTGCGCGCGCAGCTTCAATTTGCGCAACAGCGATCTTACCAAGAGCTCTTCTCCAGA 3377
DB 1081 TGTGCGCGCGCAGCTTCAATTTGCGCAACAGCGATCTTACCAAGAGCTCTTCTCCAGA 1140
QY 3378 GGGGAGTGTGGGGATTAATCAATCTGATTAAGTGTGAACTGTCCAGATGACATCCG 3437
DB 1141 GGGGAGTGTGGGGATTAATCAATCTGATTAAGTGTGAACTGTCCAGATGACATCCG 1200
QY 3438 AGGCAATTTTCCCTCGAGTGGATGATGTCCTGTTGTCTGAGATTCGAGATGATGATG 3497
DB 1201 AGGCAATTTTCCCTCGAGTGGATGATGTCCTGTTGTCTGAGATTCGAGATGATGATG 1260
QY 3498 TCGCTAGCTGAGTAAGAAACCTGCGAATTAAGTATTCGCGAGATTAATACGCAATTGA 3557
DB 1261 TCGCTAGCTGAGTAAGAAACCTGCGAATTAAGTATTCGCGAGATTAATACGCAATTGA 1320
QY 3558 GATTAAGTGTGATGATGACAG 3579
DB 1321 GATTAAGTGTGATGATGACAG 1342

RESULT 5
US-09-738-626-2905
Sequence 2905, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAMA, SATOSHI
APPLICANT: ANDO, SEIKO
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162


```

FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1264)
OTHER INFORMATION: RXN00043
US-10-781-014-413

Query Match      20.7%; Score 1237.4; DB 18; Length 1287;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1256; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 242 AACAGAGGCTTCAGTCCGAGATTAATTAACCTAGATCCGTAAGACATATAC 301
DB 1 AACAGAGGCTTCAGTCCGAGATTAATTAACCTAGATCCGTAAGACATATAC 60
QY 302 GTTCCTATGCTTGTGAGAGAACCAATTAACCTAGAGATGCGAAGTGTGATTA 361
DB 61 GTTCCTATGCTTGTGAGAGAACCAATTAACCTAGAGATGCGAAGTGTGATTA 120
QY 362 TCAAGAAAATGCGAGTCAAGAGTAAATAATTTGAGGAGAAATTTGCCCCCTCGG 421
DB 121 TCAAGAAAATGCGAGTCAAGAGTAAATAATTTGAGGAGAAATTTGCCCCCTCGG 180
QY 422 GGTGATGATGAGCTTTCTCAACTCGAAAACGCGATCATCAAGAACTCTGAGAAC 481
DB 181 GGTGATGATGAGCTTTCTCAACTCGAAAACGCGATCATCAAGAACTCTGAGAAC 240
QY 482 AGCAGCTTAAAAACGCGATTCACCCCGAACTCCCGAGATTGTTCCGGTTTATGA 541
DB 241 AGCAGCTTAAAAACGCGATTCACCCCGAACTCCCGAGATTGTTCCCGTTTATGA 300
QY 542 TCTTCATTAATCAGGGGAAAGGGTGGCGGTTTCTCAAGGGAACCGAGACAGCGAG 601
DB 301 TCTTCATTAATCAGGGGAAAGGGTGGCGGTTTCTCAAGGGAACCGAGACAGCGAG 360
QY 602 GAACACCGGCGAGTATCACCGCAAGATGCGACGACGATGATGTCAGACATGCTTC 661
DB 361 GAATGCGGCGAGTATCACCGCAAGATGCGACGACGATGATGTCAGACATGCTTC 420
QY 662 GGGCGGCGTGAAGCAGCTGCGAGCGAGTGGAAAACCTTATTCCTTGTGTGAAGT 721
DB 421 GGGCGGCGTGAAGCAGCTGCGAGCGAGTGGAAAACCTTATTCCTTGTGTGAAGT 480
QY 722 CCGCTGTGCGGATTCACCTGAGGGCCCTTTCATCAAGCATGCGGTGTGCTCA 781
DB 481 CCGCTGTGCGGATTCACCTGAGGGCCCTTTCATCAAGCATGCGGTGTGCTCA 540
QY 782 AAACCGGATTTATTTTCCCGGCAACCAAGATCTTCCCGGGTGAATCCATGCGG 841
DB 541 AAACCGGATTTATTTTCCCGGCAACCAAGATCTTCCCGGGTGAATCCATGCGG 600
QY 842 AAAAGTTGATCAATCGATCAAGTAGCGCGGAAATGACATCTTTCTGAGCTTCT 901
DB 601 AAAAGTTGATCAATCGATCAAGTAGCGCGGAAATGACATCTTTCTGAGCTTCT 660
QY 902 CGATCTCTGCGAGCGGACACATCATTTGCTTCTTCCGGGACACATGATCAGATTGA 961
DB 661 CGATCTCTGCGAGCGGACACATCATTTGCTTCTTCCGGGACACATGATCAGATTGA 720
QY 962 TACCACTACGAGCAATGCTTGGCTAAGAAAAAATGATACGGTCAAGCTACGCA 1021
DB 721 TACCACTACGAGCAATGCTTGGCTAAGAAAAAATGATACGGTCAAGCTACGCA 780
QY 1022 TTTGTTCAATGCGATGCTTCGCTGATCATAGGGCTCCCGGACGCGTGGCGCTTGTCT 1081
DB 781 TTTGTTCAATGCGATGCTTCGCTGATCATAGGGATCCCGGACGCGTGGCGCTTGTCT 840
QY 1082 TGGTGGGCGACGTCGCGGAGACGATATGTTGATGATGCGGACGCGTGAATTTGC 1141
DB 841 TGGTGGGCGACGTCGCGGAGACGATATGTTGATGATGCGGACGCGTGAATTTGC 900
QY 1142 CGATGAAACGGTGAATGCTGTTTCAACAAAGCTTTTTCATCAAGACCGCATGA 1201
DB 901 CGATGAAACGGTGAATGCTGTTTCAACAAAGCTTTTTCATCAAGACCGCATGA 960

```

```

QY 1202 AGCCGCGGAATGCCAGACGGGATACATTTTGGGCGTTTGAAGTCAACCGCA 1261
DB 961 AGCCGCGGAATGCCAGACGGGATACATTTTGGGCGTTTGAAGTCAACCGCA 1020
QY 1262 TGAAGTCGCCGCTGTGCGGATGCGCGCCATGCGCCGGGGGACACGACACTAGCGAG 1321
DB 1021 TGGCGTCGCCGCTGTGCGGATGCGCGCCATGCGCCGGGGGATACGACACTAGCGAG 1080
QY 1322 TGAAGTCGTCACACGTCGCGAGGGGTATGACGCTTATCGACGGAACCTTCAACCTTC 1381
DB 1081 TGAAGTCGTCACACGTCGCGAGGGGTATGACGCTTATCGACGGAACCTTCAACCTTC 1140
QY 1382 AACCGTCGCCGCTTAAATTCGAGTCTTGGGATGACGAAATCGCTAAATCAACCTTCG 1441
DB 1141 AACCGTCGCCGCTTAAATTCGAGTCTTGGGATGACGAAATCGCTAAATCAACCTTCG 1200
QY 1442 AAATTTGTGCTTGTGACTCAAAACGCGGAGTGCAAAAGTCAATTAAGTCAATCACT 1501
DB 1201 AAATTTGTGCTTGTGACTCAAAACGCGGAGTGCAAAAGTCAATTAAGTCAATCACT 1260
QY 1502 ACTTTAGTACGATTAATCTATCTTG 1528
DB 1261 AATTTAATACGAGCAAAATCTTCTTG 1287

RESULT 7
US-10-781-014-415
Sequence 415, Application US/10781014
Publication No. US20040180408A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Habermeyer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
FILE REFERENCE: BGI-126CPN
CURRENT APPLICATION NUMBER: US/10/781,014
CURRENT FILING DATE: 2004-02-17
PRIOR APPLICATION NUMBER: US 09/602,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/143,208
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 784
SEQ ID NO 415
LENGTH: 1287
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1264)
OTHER INFORMATION: RXN00043
US-10-781-014-415

```

Query Match 20.7%; Score 1237.4; DB 18; Length 1287;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1256; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

242 AACAGAGGCTCAAGTCCGAGATATTAACCTAGATCCGTAGACATATGACATATC 301
1 AACAGAGGCTCAAGTCCGAGATATTAACCTAGATCCGTAGACATATGACATATC 60
302 GTCTATGCTTCTGAGAGAAACCAATTAACCTAGAGAGAGAGAGAGAGAGAG 361
61 GTCTATGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
362 TCAAG 421
121 TCAAG 180
422 GGTGATGATGAGCTTCTCAACTCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 481
181 GGTGATGATGAGCTTCTCAACTCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 240
482 AGCAGCTAATAAG 541
241 AGCAGCTAATAAG 300
542 TCTTATATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
301 TCTTATATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
602 GAACACCGGAGAGATCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
361 GAATGCGGAGAGATCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
662 GCGCGCGGAG 721
421 GCGCGCGGAG 480
722 CCGTCTGAGAGAGATCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
481 CCGTCTGAGAGAGATCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
782 AAACCCGAGATTCATTTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
841 AAACCCGAGATTCATTTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
842 AAAGGTTGATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 901
601 AAAGGTTGATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 660
902 CGATCTGAG 961
661 CGATCTGAG 720
962 TACCACTACAG 1021
721 TACCACTACAG 780
1022 TTTGTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
781 TTTGTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
1082 TGTCTGAG 1141
841 TGTCTGAG 900
1142 CGATGAG 1201
901 CGATGAG 960
1202 AGCGCGGAG 1261
961 AGCGCGGAG 1020

1262 TGGAGTCCGCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1321
1021 TGGAGTCCGCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1322 TCAGTTCGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1381
1081 TCAGTTCGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
1382 AACGTCGCGGATTAATTCGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1441
1141 AACGTCGCGGATTAATTCGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1442 AAATTTGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
1201 AAATTTGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1502 ACTTAACTAGAGATTAATTCGAGTCTGAGAGAGAGAGAGAGAGAGAGAG 1528
1261 AATTTAAATAG 1287

RESULT 8
US-09-738-626-2907
; Sequence 2907, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2907
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-2907

Query Match 18.6%; Score 1110.4; DB 9; Length 1152;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

354 GTGCAATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 413
1 GTGCAATATCAAG 60
414 CCCCTCGGAGAGATGAGTCTTCTCAACTCGAAGAGAGAGAGAGAGAGAGAG 473
61 CCCCTCGGAGAGATGAGTCTTCTCAACTCGAAGAGAGAGAGAGAGAGAGAG 120
474 GAGAGACAG 533
121 GAGAGACAG 180
534 TTTATATGATCTTCAATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 593
181 TTTATATGATCTTCAATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240


```
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1513)
; OTHER INFORMATION: RXA00314
US-10-494-541-9

Query Match
Best Local Similarity 100.0%; Pred. No. 8.4e-79;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5674 ACCGTAAACACTAGGTTTGGGCGAAATGAAAGCGGCGCAACCTGCTCAACGTCG 5733
DB 1 ACCGTAAACACTAGGTTTGGGCGAAATGAAAGCGGCGCAACCTGCTCAACGTCG 60
QY 5734 AAAGAAAGAGCGGTGCGACCAACACATTAAGTTGAAACCTTAGTGTTCGACACAGT 5793
DB 61 AAAGAAAGAGCGGTGCGACCAACACATTAAGTTGAAACCTTAGTGTTCGACACAGT 120
QY 5794 TAGACTAGGGAGAGTGAAGTCTTACGCACTTTTGAACCCGGTACCCGTTGAGATT 5853
DB 121 TAGACTAGGGAGAGTGAAGTCTTACGCACTTTTGAACCCGGTACCCGTTGAGATT 180
QY 5854 TAAACCTGTTCAACGAGTGCCTCGGTTACCTGTGTTGTTGTTGTTGTTGTTGTTG 5913
DB 181 TAAACCTGTTCAACGAGTGCCTCGGTTACCTGTGTTGTTGTTGTTGTTGTTGTTG 240
QY 5914 ACCCCACATTGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5967
DB 241 ACCCCACATTGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 294

RESULT 12
US-09-746-660A-31/C
; Sequence 31, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 31
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (101)..(1036)
; OTHER INFORMATION: RXA00044
US-09-746-660A-31

Query Match
Best Local Similarity 96.6%; Score 215.2; DB 10; Length 1059;
Matches 220; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AGTCGTCAGACGCGACCATTTGATGTGTGTCACCGAGCTTGCGAGGCTTTTCAATCT 60
DB 228 AGTCGTCAGACGCGACCATTTGATGTGTGTCACCGAGCTTGCGAGGCTTTTCAATCC 169
QY 61 ACCCTCCCGTGGCGGTGAGTGGGTTCATTACGGGTGGATACGCGCGTGAAGTTGCG 120
DB 168 ACACCTCCCGTGGCGGTGAGTGGGTTCATTACGGGTGGATACGCGCGTGAAGTTGCG 109
QY 121 GAACCATGTTGTTCTTGTGGTTGAGGAAAGAGTGGCGGTGAGAAAGTTTCAAGTG 180
DB 108 GAACCATGTTGTTCTTGTGGTTGAGGAAAGAGTGGCGGTGAGAAAGTTTCAAGTG 49
QY 181 TCTGCAGTTTTTAAGTTATGATCATCATGCTTGGAAGCTGAGGTAT 228
DB 48 TCTGCAGTTTTTAAGTTATGATCATCATGCTTGGAAGCTGAGGTAT 1

RESULT 13
US-10-282-122A-17961
; Sequence 17961, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17961
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Corynebacterium diptheriae
```


US-10-282-122A-17961

Query Match 3.5%; Score 208; DB 17; Length 777;
 Best Local Similarity 57.8%; Pred. No. 1.9e-52;
 Matches 428; Conservative 0; Mismatches 305; Indels 7; Gaps 3;

QY 1548 ATGGAATCACTTCTGCAAGCGAGCAAGAGTGGCAAGAGTTCGATCTTCAATC 1607
 DB 1 ATGGAATCGTAATCAACCCCAAGCAAGAGCGCGCTGTATCCGCCGCACTCTTA 60
 QY 1608 GCACCTTGGCCCAAGAGGTGGAACCTTGGGGCTTGCAACAGATCTCACTAGAT 1667
 DB 61 GAAGAATAGCAGTGAAGGAAGACCTGGGCTTGCTACCGATCTACCTCGAGGC 120
 QY 1668 ACTTACCAAGAGTCTATGGCATGTATGAAGCTGGGAAAGTGTATTCAGAACTGCAAG 1727
 DB 121 ACCTACCAAGAGTCTATGGCATGTATGAAGCTGGGAAAGTGTATTCAGAAAGTCCAA 180
 QY 1728 GCATCTTGTGGATATACGTTGGGACTTAACCCGATGAGTGAAGAAAGCTACTTAA 1787
 DB 181 GCATTTACTCTGACGAATATCGTGGATTCCTCGGAAACAGAAAGCTACTACAGC 240
 QY 1788 ACCATTCGAAAGAGTCTACTGACATGACATGATGTTGAGAGGTTTACAGCCCA 1847
 DB 241 ACCATTCGACGGGAATTTACTTCCATATCGAATATCCAGAGAGAGGTTTCAACCT 300
 QY 1848 GATGTGCAAAACCTTATCTATCGAAGCAAGCTGCAAGATGAGGCAAAATCGCTGA 1907
 DB 301 GACGGCAGCGCAAGAACCCAGGACAGCAAGAAATAGCATGCTGATCGTAGAA 360
 QY 1908 GAATCC---GTTGAAGTTCAAAATCCTTGGCATCGGGGAAAGG--CACATGCTTTCANT 1963
 DB 361 AAAGCGGTGTTGACATTCAGATTTCTTGATTCGAACCGATGTCATTCCTTTTAAAC 420
 QY 1964 GAACCATCATCTTCTCTGTACAGACTGACAAAGTTCAGAGCCCTGACCCCTTAAACTGTG 2023
 DB 421 GAGCCAACTCGCTATGCGAAGCCGTAACGCTATCAAGACCTGACCCCAACAGCGTA 480
 QY 2024 GAGGCAAGCTCGATTTCTTCAACCATCGA---AGAGTCCCAACCCAGCCGCTACG 2080
 DB 481 CGCGAATATTCGATTTCTTAAATGCAACGATCAAGATTCACACACATGATGACC 540
 QY 2081 CAGGTTTGGGCACTTTGTCCCGCGCAAAACATGTTGTTGGCAACTGGTGAAGGA 2140
 DB 541 CAAGGAATCGGCACTATTCGCGAGGCCCAACCTGCTGATCTCTGCTTTGGCAGAAAT 600
 QY 2141 AAAGCCGACGCAATCCGCGAACTGTGAAGGCCCAAGTACTGCTTTTGGCCAGGTTCC 2200
 DB 601 AAAGCAGATGCAAGTCAAGCCATGTGTGAAGGCCCAAGTCAAGCCGTGCCCACATCT 660
 QY 2201 ATCTGTAGATGCAACAACATGCAACCATGATGTTGATGAGAGCAAGTATCCAACTG 2260
 DB 661 GTTTTGACACTGACAGATGCACTGTGATTTGTGATGAGGCGCAGCAAAACTT 720
 QY 2261 GAAAAAGCTGATCACTACCG 2280
 DB 721 GAGCACTCACTACTACCG 740

RESULT 14

US-10-282-122A-17959
 ; Sequence 17959, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zykend, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282.122A
 PRIOR FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 17959
 LENGTH: 1212
 TYPE: DNA
 ORGANISM: *Corynebacterium diptheriae*
 US-10-282-122A-17959

Query Match 3.2%; Score 192.6; DB 17; Length 1212;
 Best Local Similarity 53.5%; Pred. No. 1.4e-47;
 Matches 433; Conservative 0; Mismatches 364; Indels 12; Gaps 1;

QY 511 AACTCCCAAGATTTTCCGGTTTATGATCTTCAATCAAGTGAAGAGTGGCG 570
 DB 194 ACCACCCCTGTGCTGCAAGTCTTGCCGACATCCAAATCAAGTGAAGAGGAGT 253
 QY 571 CTTTCTTCAAGGAAACGAGACCAAGGAGAAACCGCGCATATCACCGCAATG 630
 DB 254 CTTTCCCAACTCCGACTACGACGAGATGCGATGCGCGACACACCGCGCAATG 313
 QY 631 GCAGACCGTATGTTGCAAGCATGTTTCCGGCGCGGCTGACGCACTGGCAAGCGAG 690
 DB 314 GCTGACACGCTCTGCGCTGCGACATATCATGCCCCGAGACACCTTACTGCCGAAAC 373
 QY 691 TGGAAACCTTATTCCTTGTGTGAAAGGTCTGTGCGGATTCACCTCGAGGCGC 750
 DB 374 TATCCCTCTTACGAGATCTTCCGACGCCGCTGAAATCAAGCAATCCAGCGGAAGCC 433
 QY 751 CTTTCAATCAAGCATGCGCTTGTGTCTCAAAACCGGATTTATTTTCCGGCAAC 810
 DB 434 CTTTGTAAACCCGTGCGCTGCGCGCACAAAGCCCAAGAACGATATCTTAGGCGACC 493
 QY 811 CAACAGATCTTGGCCGGGTGATCCATGCGGGAAGAGTGGATCAATTCATGATCACTAG 870
 DB 494 CAGAACTGTTCAAGAAATATGTCAGAGCGGACCGGCTGCTCAATTCATGACTTCC 553
 QY 871 CGCCGAAACTGACATCTTGTAGCTTCTGATCTCTGCGACGACCAACATCATTTG 930
 DB 554 CCCCAGAAACGACACAGCAAGGAATCATGACCTCTGCGCGGAAACAAATCATTCG 613
 QY 931 CTTTCTTGGGACATGATGCAATTTTATATACATACACAGGCAATTCCTTGGCTA 990
 DB 614 TCTTCTTGGACACACGACGAGACTTCTCACTGATCCGAACAGCCCTTCTTACGAG 673
 QY 991 AAGAGAAATATGACGCTACCGCTACGATTTGTTCATGAGATGCTCCGCTGATC 1050

```

Db      674 TAGGACCCGGTGCACACTGTGACCCGCAACCCACCTATTATATGCAATGCGACAAATACAC 733
Qy      1051 ATAGGGCTCCCGGACAGCGTGGCGCTTCTGCTGCGGACAGTCCGGGGAGCATATG 1110
Db      734 ACCGTGACCCCGGTGACGCGGACGCTTATCGACCCGACGCTGAGGCAACGCCACAG 793
Qy      1111 TTGAGTGTATGCGCGGACGCGGTGATTTGGCCGATGGAACG-----TCGATC 1158
Db      794 TAGAACTAGTCCGCGATGCGATCACCCTGATGACCAACAGTGGCGCATGTGATCGACT 853
Qy      1159 TAGCTGCTTCAACAACAGCGCTTTTCATACAGGACCCGATGGAAGCCGCGGAAATGCCAG 1218
Db      854 CCGTAGTGGCGGATCGAGTAGAGCTTGTCTCCGACCCGATGGGCGGAGAGAAAGAG 913
Qy      1219 ACGGTGAGTACATTTTGGCGCTTTTGAAGTCACCGTCAACCGATGAGTGGCCCTGTCG 1278
Db      914 ACGGGGACTACCTTCTCGGAGCCCTGCGCGTCAACGTAAGACTCCGTCGACGACTGA 973
Qy      1279 GGGATGGCGGCGCATCGCGGGGGGACG 1307
Db      974 CACCAACGACGCGAGGAGGCGTATC 1002

RESULT 15
US-09-974-300-829
; Sequence 829, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 829
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-829

Query Match      3.2%; Score 192; DB 9; Length 975;
Best Local Similarity 51.4%; Pred. No. 1.9e-47;
Matches 472; Conservative 0; Mismatches 440; Indels 6; Gaps 1;

Qy      4240 TGTTCGTGATGATCAACCTGATGCGATGCGCGCTTCTTGCAGTGTGCT 4299
Db      27 TCTCGCTAATATGATTAATCTTATGCGGAGACTGCTTTCGATTTCTCCGCGCTAAT 86
Qy      4300 TGTTCACCGCAACCAACCGTTTGGTGGCAATGATGCTTGGGCGCGCGCAATGGTAT 4359
Db      87 CCGCTGTGTGCGCGCTCAACCGTTGGCGGAAATCCGCTGCTCGGCATGTCTCGGCGCT 146
Qy      4360 GCGGATGTGTGCCCAACCTGTTAAGGCTTACGAGCTAGCGGCGGACCATGACCGCGGG 4419
Db      147 CATGCTTGTGACCCCTGATTTGTTAAACCCCTGGGGATACGGGGCTGCGAGCAAAAGCGG 206
Qy      4420 CGAAATGCAATGTGCTCCCTGTTGGTGTGATGTTGCTCAAGCTGTACCAAGGCGAC 4479
Db      207 AGAGATTCCTGTCTGGAATTTATGCTTTGGAAGTGAAGAGGTGCGCTACCAAGGGGCA 266
Qy      4480 CGTGTCTCTGTGCTGTGTGCTCTTGTGATTTCTGCGCAAGATGAGAAATTCCTGACAA 4539
Db      267 GGTGTCTTCCATTTTGTGCTGTCTTATTTGTGCGCAAAATTTGAGCGCTTTTGTGCAA 326
Qy      4540 GCGACTCATGGGCACTGACAGACTTCTGATCAACCCAGTGTGACTGTGCTACCGG 4599

```

```

Db      327 GCGAGCGCTGAGACATTCACCTCTCTGTTGGGACCCGATTAACGCTCTTGTACAGG 386
Qy      4600 TTTCCTTACGTTCAATGCTTATTTGTCACGAATGCGGTGGGTGACTTCTGGCACA 4659
Db      387 TTTCCTTCAATTATTTGCTTGTGACCGATTAACATTTGCCATCGGAAATGTTTAAACATC 446
Qy      4660 CCGTGTGACGAGACTATATGATTTGGGTGTGTCATGTCGCGGTGCTTTTGGCTGTGT 4719
Db      447 AGGCTTATTTGTGTCTTGAACAAATGCGGTACTGAGCGGACGTTATATGAGAGGCT 506
Qy      4720 TTACTACCAATTCGTTATCACTGTGTCACCAAGCTCTTCCGCAATTTGAGCTGAGCT 4779
Db      507 TTAACCGCTCTGCGCATTAACCGAATGCAATATTCATTCTTGGCGTTGACCTTACACT 566
Qy      4780 GTTC-----AACGAGGTGATCTTCACTTTCGCAACCGCATCATGCGCAATATGCG 4833
Db      567 GATCGGCTCAAAAGCTCGGCGGAAACATTTTATAGCGGATGCTCGGCTGTCCAATATTGC 626
Qy      4834 GAGAGTGCAGATGTTGGCAGTGTCTTCTAGCGAAGATGAAAGCTCAAGGCGCT 4893
Db      627 ACMAAGTTCAAGCGCGCTTGGATGATGTTATTTGCAAGATGAGAAAGCAAGAAAGGCT 686
Qy      4894 TSCAGTGTCTCAGGTGTCTCGCTGTTCTTGTATTAACAAGCTGCGATCTTGGTGT 4953
Db      687 GTTCCTGACATCCGAATTTGACTTATCTGGAATTAACGAACCGCATTTGTCGAGT 746
Qy      4954 GAACCTGCGCTGCGCTGCGCTTCTACATTTGATTCGATACCGGACTATGCTGGCGC 5013
Db      747 GATCTCAGATACAGATTTCCGTTGCTGATGCGCATGATCAAGTTGCGGGAATGCGGAT 806
Qy      5014 TTTGATGACCTTTGATATCAAGGCAAGTTCGTTGGCGCTGCAAGTTTCTTGGGTGT 5073
Db      807 GTTCAATTTCTCGCAAGGGGTTTGGCAAGCTGTGCGCGCTGCGCGGCTAACCTGGAAT 866
Qy      5074 TGTTCCTATGATGCTCCAGATATGATGATGTTCTTGTGTTGGCGGAGTATTAACCTTGT 5133
Db      867 TTTCTCTATTAAGACCAATCTGAGGAGCGTTTGCATCGGAATGCGATGCTGTGAT 926
Qy      5134 CATGCAATTCGCGGACG 5151
Db      927 CCGCGCGTTTGGCGGAGC 944

```

Search completed: March 9, 2005, 20:15:11
Job time : 2997 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 19:13:49 ; Search time 16576 Seconds
(without alignments)
13706.903 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969

Sequence: 1 agtcgcgtgcagccaccatc.....tgatatttcgcgcgtgaa 5969

Scoring table:

IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	174.8	2.9	917	8	BZ685850 PUBER66TD
2	132.4	2.2	770	7	CN823265 OA.spibn
C 3	123.6	2.1	890	6	CD499281 CDA38-D12
4	113.4	2.1	1030	6	CD499280 CDA38-D12
5	113.4	1.9	745	5	BP705177 BP705177
6	112	1.9	732	7	CK807155 AGENCOURT
7	111.8	1.9	760	7	CK805715 AGENCOURT
8	105.4	1.8	899	5	EX714658 BX714658
9	104.4	1.7	811	7	CF346432 AGENCOURT
10	104.4	1.7	896	7	CF592490 AGENCOURT
11	104.4	1.7	1177	3	CR696680 Tetradon
12	104.4	1.7	1178	3	CR690088 Tetradon
13	104.4	1.7	1206	3	CR701935 Tetradon
14	103.8	1.7	818	4	BM017859 BM017859
15	103.2	1.7	662	4	AL656223 AL656223
16	103.2	1.7	902	7	CF378888 AGENCOURT
17	102.8	1.7	1173	3	CR703987 Tetradon
18	102.8	1.7	1160	3	CR699192 Tetradon
19	102.8	1.7	1193	3	CR698198 Tetradon
20	102.8	1.7	1203	3	CR674049 Tetradon
21	102.8	1.7	1226	3	CR669603 Tetradon
22	102.6	1.7	1242	3	CR665914 Tetradon
23	102.6	1.7	818	5	BX458246 BX458246
24	102.6	1.7	846	7	COS72502 AGENCOURT

25	102.6	1.7	870	9	AY416649 Homo sapi
26	102.6	1.7	832	5	BX374555 BX374555
27	102.6	1.7	1033	1	AL560879 AL560879
28	102.6	1.7	1049	5	BX406188 BX406188
29	102.6	1.7	1051	4	BM476096 BM476096
30	102.6	1.7	1063	5	BM910583 AGENCOURT
31	102.6	1.7	1114	3	BX421863 BX421863
32	102.6	1.7	2229	3	CR609510 full-1eng
33	102.4	1.7	1040	9	AL152142 Anopheles
34	102.2	1.7	669	2	BE790435 BE790435
35	102.2	1.7	847	5	BU171522 AGENCOURT
36	101.8	1.7	1003	1	AL546345 AL546345
37	101.4	1.7	949	5	BX324827 BX324827
38	101.2	1.7	892	3	CR713846 Tetradon
39	101.2	1.7	917	5	BX428009 BX428009
40	100.2	1.7	875	5	BU158666 AGENCOURT
41	100	1.7	746	4	BI758774 BI758774
42	100	1.7	812	4	BI825295 BI825295
43	100	1.7	905	5	BU179037 AGENCOURT
44	100	1.7	934	5	BO929688 AGENCOURT
45	100	1.7	984	5	BU167870 AGENCOURT

ALIGNMENTS

RESULT 1
BZ685850 917 bp DNA linear GSS 05-FEB-2003
LOCUS PUBER66TD ZM.0.6.1.0 KB Zea mays genomic clone ZMBETA035K11,
DEFINITION genomic survey sequence.
ACCESSION BZ685850
VERSION BZ685850.1 GI:28245462
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 917)

REFERENCE
AUTHORS Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Contact: Cathy Whitefaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org.
Seq primer: TP
Class: sheared ends.

FEATURES
source
1..917
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBETA035K11"
/clone_lib="ZM.0.6.1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN

Query Match 2.9%; Score 174.8; DB 8; Length 917;
Best local Similarity 52.7%; Pred. No. 3.4e-39;
Matches 483; Conservative 0; Mismatches 412; Indels 21; Gaps 4;

QY 3930 TGAAGGACCTTGAACCTGCGCGATGTCATGTCATCGGCGCCAGCGATGTGG 3989
DB 914 TCAAGGTTCTTCTTCCACCGGCGGCTGTCCAGGTGTAATCGCCCGGTAAGTGG 855

Db 695 GGGCCCATGCTGGGGGTCCTTTTCGGCTCTGTACGACCCCTGATATCCCG 754

Qy 4744 TCTGACCAATCCTTC 4759

Db 755 CATGCACCAATGTTT 770

RESULT 3
CD499281 890 bp mRNA linear EST 12-JUN-2003
LOCUS CDA38-D12.3, SHGC-CDA Gasterosteus aculeatus cDNA clone
DEFINITION CDA38-D12.3, mRNA sequence.
ACCESSION CD499281
VERSION CD499281.1 GI:31426312
KEYWORDS EST.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 890)
REFERENCE Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J. and Myers, R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
CONTACT: Kingsley, DM
HIMI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cmgm.stanford.edu
Plate: 38
High quality sequence start: 13
High quality sequence stop: 799.
FEATURES
source
1..890
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas river, CA"
/db_xref="taxon:69293"
/clone="CDA38-D12"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
/note="Vector: lambda ZAP Express/PBK-CMV. Site 1: EcoRI
(5' adaptor); Site 2: XhoI (3' linker primer). The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dt sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of PBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual PBK-CMV phagemid clones for
EST sequencing."

ORIGIN
Query Match 2.1%; Score 123.6; DB 6; Length 890;
Best Local Similarity 52.3%; Pred. No. 4.7e-24;
Matches 346; Conservative 0; Mismatches 309; Indels 7; Gaps 3;

Qy 1606 TGGACACCTTGGCCACAAGGGTGAACCTTGGGGCTTGCACAAGATCTTCAACCACTGA 1665

Db 771 TCCGACCTGTGTGGCGACAGATTTTTCACCTCTGGGGCTTCCACAGAGACACCCCAATGG 712

Qy 1666 GAACTACCAAGAGCTCATTCGATGTATGAAGCTGGGAGAGTGTATTCAGAAACTGA 1725

Db 711 GTTTTACAGAACTAATGATCTACTACAAATGGCCAAATTCATTCTAGTATGA 652

Qy 1726 AGGCATTTCTTGTGATGAATACGTGGGACTAACCCGTGACGATGAAAACAGTACTTTA 1785

Db 651 AAACCTTGAACATGATGAATAGTAGAGACTTCCAGAGATACCCGTGAGATCACT 592

Qy 1786 AAACCATTCGCAAGAATTCAGTACACCATGACATCGTTGATGAAAGGCTTACAGCC 1845

Db 591 CCTTCATGTGGAATCAATTCCTTCAAGCATGTAGACATTAACAGAGAACTCATCATCC 532

Qy 1846 CAGATGTGCAAAACCTGATTCATGCAAGAGAGCTGCAAGATGATGAGGCAAGAT---CG 1902

Db 531 TTGATGGCAAGCCCGCTGACTTCAAGAGAGTGTGACGATTTGAGGAAAGATACAG 472

Qy 1903 CTGCAGATCCCTTGAAGTTCAATCTTGGCATCGCGGAAACGG-CACATGCTTTCA 1961

Db 471 CTGCGGAGGAGATCGAGCTTTGTGAGAGTATGACACAGATGGCCACATTTGCTTCA 412

Qy 1962 TTGAACCATCATCTTCTCTGTGAGACTGACAAAGTCCAGGCGCTGCACCTTAACCTG 2021

Db 411 ATGAGCTGTGTTCAAGCCTGTGTTCCAGAGACTTAGGTTGAAAGCCTGGCAAGACACTA 352

Qy 2022 TGAGAGCAAGCGTGAATCTTCAA---CACATGAAAGAGTCCCAACCCAGCGCTCA 2078

Db 351 TGTGACCAATCTGTGATTTCTTGTATGAGATCTTCAAAAGTCCGACATGACCTGA 292

Qy 2079 CCCAGGTTTGGGCACTTGTCCCGCGCAAAACATCGTGTGTGGTGGCACTGTGAAG 2138

Db 291 CGGTGGAGTGGGCAAGTCAATGACGCAAAAGAGTCTGATTTCTATCATCTGAGAC 232

Qy 2139 GAAAAGCCGACCCATCCGCGGAATGTGGAAGGCCCACTGCTTTTGGCCAGGTT 2198

Db 231 ACAAGGCTTTTGCCTTATACAAAGCTATAGAGAAAGTGTGATCAATCATGTGACAGTGT 172

Qy 2199 CCATCCTGTGATGACAAACATGCCACATCATCGTTGGATAGAGAGATGCCAAGC 2258

Db 171 CCGCTTCCAGACGACCCGACAGACAGTTTGTGTGATGAGACGCCACATTTGAAAC 112

Qy 2259 TG 2260

Db 111 TG 110

RESULT 4
CD499280 1030 bp mRNA linear EST 12-JUN-2003
LOCUS CDA38-D12.3, SHGC-CDA Gasterosteus aculeatus cDNA clone
DEFINITION CDA38-D12.3, mRNA sequence.
ACCESSION CD499280
VERSION CD499280.1 GI:31426311
KEYWORDS EST.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1030)
REFERENCE Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J. and Myers, R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
CONTACT: Kingsley, DM
HIMI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cmgm.stanford.edu
Plate: 38
High quality sequence stop: 822.
FEATURES
Location/Qualifiers

QY 1956 CTTTCAATGAACCATCATCTTCTCTGTCAGACTGCAAAAGTCCAGCGCTGCACCTTA 2015
 Db 480 CCTTCATATGAGCCAGGGGTCAAGTCTGTCTCCAGAGCCCGAGTGAATACTTAGCCATG 539
 QY 2016 AAAGTGTGAGAGCAACGCTGCAATCTT---CAACACCATTCGAAGAGTCCCAACCCAG 2072
 Db 540 ATATGATTTCTGGCCATGACAGGTTCTTTGACGGCAACCTCTCCAAAGTCCCAACCATG 599
 QY 2073 CCGTCAACCAAGGTTGGGCACTTTGTCCCGCGGCAAAACATGCTGTGGTGGCACTG 2132
 Db 600 CCTGCTCTGTGGTGTGGGCACTGTATGACATCCAAAGAGTCATGATTCATCATCAG 659
 QY 2133 GTGAAGAAAGAGCCGACGTCATCCGGGGAACCTGTGAAGCCCAAGATCTCTTGGC 2192
 Db 660 GAGCCCATTAAGCTTTGCTGTGTATTAAGGCTATTGAGAGAGGTGAATCATATGTGA 719
 QY 2193 CAGTTTCATCTCTGTATGATGAC 2215
 Db 720 CAGTTCTGTCTTCCAGACGAC 742
 RESULT 6
 CK807155 732 bp mRNA linear EST 26-FEB-2004
 LOCUS AGENCOURT 19145816 NICHD_XGC_Te2 Xenopus laevis cDNA clone
 DEFINITION IMAGE:7208717 5', mRNA sequence.
 ACCESSION CK807155
 VERSION CK807155.1 GI:43396531
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 732)
 REFERENCE NIH-MGC http://img.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Daniela S. Gerhard, Ph.D.
 COMMENT Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Igor B. David
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.llnl.gov
 Plate: LHAM15095 row: a column: 03
 High quality sequence stop: 707.
 Location/Qualifiers
 1..732
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:7208717"
 /tissue_type="Pooled samples from 6 adult Xenopus testis"
 /lab_host="DH10B Tera"
 /clone_1lb="NICHD_XGC_Te2"
 /note="Organ: testis; Vector: pExpress-1; Site_1: EcorV;
 Site_2: NotI; RNA obtained from 6 adult male testis. cDNA
 was primed using oligo-dT primer:
 5'-pGACTGTTCTAGATCGCAGCGCGCCGCT(125-3'- and cloned into
 the EcorV/NotI sites of pExpress-1. Size-selection >1kb
 resulted in an average insert size of 1.25 kb. This is a
 primary library (normalized primary library is
 NICHD_XGC_Te2N) and was constructed by Express Genomics
 (Frederick, MD). Note: this is an NIH_XGC library"
 ORIGIN
 Query Match 1.9%; Score 112; DB 7; Length 732;

Best Local Similarity 54.0%; Pred. No. 1,2e-20;
 Matches 296; Conservative 0; Mismatches 245; Indels 7; Gaps 3;
 QY 1611 CCTTCGCAACCAAGGTTGGAACCTTGGGCTTGGCAACAGATTCCTACCACTGAGTACC 1670
 Db 135 CCGACCGCTGATAGTACTTACTTTGGGCTTCCGACAGGAAGCAACCCCTAGAGATGC 194
 QY 1671 TACCAAGACTCATTCGATGATGAGCTGGGGAAGTGTCAATTCAAGACTGCAAGCA 1730
 Db 195 TACAAAACCTAATTGAGTACCATTAATAATGAGATCTGTCTTCAAGTATGAAAACA 254
 QY 1731 TTTCTTGGTGAATGATAGCTGGGACTTAACCCGTGCGATGAAGAAACGCTACTTAAAC 1790
 Db 255 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 314
 QY 1791 ATTGCAAGAGTTCATGACATGACCAATGCAATGATGTTGTAAGAGTCTACAGCCAGAT 1850
 Db 315 ATGTGGAACAATTCTTCAAGCAATGACATGACGCTGAGAACGCCACATCTTGGAT 374
 QY 1851 GGTGCAAACTGTATTCATACGAAGCACTGCAAGATGATGAGCAAAAGT---CGCTGCA 1907
 Db 375 GGAACCGCTGAAGACCTTCAGGCGCAGATGTGACCTGTTGAAGAAAGATTGCGGCTGCT 434
 QY 1908 GAATCCGTTGAAGTTCAAATCCTTGGCAATCGG-CCGAACCGGCACATGCTTCAATTGA 1966
 Db 435 GAGGAGATTGAACCTGTTGTGTGAGGATGCTGCTGATGAGCCACATGCTTCAATGAG 494
 QY 1967 CCATCATCTTCTCTGTGCAAGACTGACAAAGTCTCAGCGCTCACCCTAACTGTGAG 2026
 Db 495 CCAAGGTCAAGTCTGTGTCTCCAGAGCCGAGTGAAGAAAGTGTGCAATGACATTTCTG 554
 QY 2027 GACAAAGCTGATTTCTTCAA---CACCATGAAAGTCTCCAAACCAAGCCGTCAACCG 2083
 Db 555 GCTAATGACAGCTTCTTGTGATGAGCAACCTCTCAAGTCCAAATGCTCCGACTGTG 614
 QY 2084 GGTGTGGGACCTTGTCCCGCGGCAAAACATGCTGTGTGAGCAATGCTGTAAGAGAAA 2143
 Db 615 GGTGTGGGACCTTGTATGAGATCCAAAGAGATGATTCATCATCGGAGCCCATTA 674
 QY 2144 GCCGACGC 2151
 Db 675 GCCTTTGC 682
 RESULT 7
 CK805715 760 bp mRNA linear EST 26-FEB-2004
 LOCUS AGENCOURT 19145650 NICHD_XGC_Te2 Xenopus laevis cDNA clone
 DEFINITION IMAGE:7210866 5', mRNA sequence.
 ACCESSION CK805715
 VERSION CK805715.1 GI:43392452
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 760)
 REFERENCE NIH-MGC http://img.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Daniela S. Gerhard, Ph.D.
 COMMENT Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Igor B. David
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.llnl.gov

Plate: L1AM15100 row: j column: 16
High quality sequence stop: 667.
Location/Qualifiers

FEATURES
source

1. 760
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:7210866"
/tissue_type="Pooled samples from 6 adult Xenopus testis"
/lab_host="DH10B Tona"
/clone.lib="NICHD_XGC_Te2"
/note="Torgan: testis; Vector: pExpress-1; Site_1: Scovy; Site_2: NotI; RNA obtained from 6 adult male testis. cDNA was primed using oligo-dt primer: 5'-pGACTAGTCTAGATCGCGAGCGGCC(17)25-3' and cloned into the Scovy/NotI sites of pExpress-1. Size-selection >1kb resulted in an average insert size of 1.25 kb. This is a primary library (normalized primary library is NICHD_XGC_Te2N) and was constructed by Express Genomics (Frederick, MD). Note: this is an NIH_XGC library"

ORIGIN

Query Match 1.9%; Score 111.8; DB 7; Length 760;
Best Local Similarity 52.4%; Pred. No. 1.4e-20;
Matches 318; Conservative 0; Mismatches 282; Indels 7; Gaps 3;

1600 TCCATATGCGACCCCTTGGCAAAGGTTGAACCTTGGGGCTTGCAACAGATCCTCAC 1659
88 TCCATATGCGACCCCTGATTAATCTTACTTGGGGCTTCCACAGAGACACGCG 147
1660 CACTGAGTACCTACCAAGAGCTCATTCGATATGAAGCTGGGGAAGTGCATTCACAGA 1719
148 CCGTAGATGCTACCAAGAACTAATGAGTACCATTAATGAAGTCTCTCTTCAAGT 207
1720 ACTGCAAGCATCTTCTTGATGATGATACGTGGGACTAACCCGTGACATGAAAAAGCT 1779
208 AATGTAACATTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
1780 ACTTTAAACATTCGCAAGAGTTCACGACCAATGCACATCGTTGATGAGAGTCT 1839
268 ATCATTCCTTCATGATGAAACATCTTTAAGACATTCATCAGAGGTGAAATGCC 327
1840 AAGAGCCAGATGCTGCAACCCGTATCATAAGCAAGCAGCTGACAGTATGAGCAAGA 1899
328 AATCTTGATGAAAGCCGACAGACCTTCAGGACAGAGTGCACCTGTTTGAATGATA 387
1900 T--CGCTCAGATCCGTTGAAGTCAATCCTTGGCATCGG-CGAAACGGCACATCG 1955
388 TTGGGCTGCTGGTGGCATTTGAACCTGTTTGGAGGTATCGGCTCTGATGTCACATAG 447
1956 CTTTCAATGACCATATCTTCTCTGACGACTGCAAAAGTCCAGGGGCTGCACCCTA 2015
448 CCTTCATAGCAGAGGTCAAGTCTGTCTCCAGAGCCGAGTGAAGAACTTAGCCATGG 507
2016 AACTGTGAGACAAGCTCGATCTT---CAACACATCGAAGAGTCCCAACCCAG 2072
508 ATACTATTCGCGCAATGACAGCTTCTTGAACGCAACTCTCCAAAGTGCACCAACATG 567
2073 CCGTCAACCAAGGTTGGGCACTTGTCCCGCGCAAAACATCGTTGGTGGCACTG 2132
568 CCTGACTGTGGGTGGGAGCTGTATGAGCTCCAAAGAGTCATGATTCATCATCAG 627
2133 GTGAAGAAAAGCCGACGCAATCCGGGGAACGTGGAAGCCGAGTGCCTTTGGC 2192
628 GAGCCCATAAAGCTTTGCTTTATTAAGGCTATTAAGAGGTGTAATCATATGTGA 687
2193 CAGGTTT 2199
688 CAGTTTC 694

RESULT 8
BX714658

LOCUS BX714658 899 bp mRNA linear EST 18-NOV-2003
DEFINITION BX714658 XGC-tadpole Xenopus tropicalis cDNA clone TTPA01516 5', mRNA sequence.
ACCESSION BX714658
VERSION BX714658
KEYWORDS BX714658.1 GI:38386145
SOURCE EST.

ORGANISM

Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

REFERENCE

1 (bases 1 to 899)
Croning M.D.R., Ashurst J.L., Taylor R., Zorn A.M. and Rogers J.
Xenopus tropicalis EST project 2001 (11_2003)

JOURNAL

COMMENT

Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE ID: TTPA01516.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Nigel Garrett.
cDNA was oligo dt primed from 5'ug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.

FEATURES

source

1. 899
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA01516"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone.lib="XGC-tadpole"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5'ug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 1.8%; Score 105.4; DB 5; Length 899;
Best Local Similarity 49.9%; Pred. No. 1.1e-18;
Matches 322; Conservative 0; Mismatches 316; Indels 7; Gaps 2;

1611 CCTTCGGCAACAGAGTGAACCTTGGGCTTGCAACAGATCCTCACCTGAGTACC 1670
111 CCACCGCTGATTAATCTTACTTATGAGGCTCCCAACAGGAAGCACCCCACTGAGATGC 170
1671 TACCAAGCTCATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1730
171 TACAAAACTTAATGAGTACCATTAAGAGAGTCTCTCTTCAATATGTAAGAA 230
1721 TCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1790
231 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 290
1791 ATTGCAAGAGTCACTGACCAATGCACATCGTTGATGAGAGTCTTAAGCCCAAT 1850
291 ATGTGAACAACCTTCTTAAGCACAATCAACAGCTGAGAAAGCCCAATCTTGAT 350
1851 GGTGAACCCCTGATCCATAGCAACAGCTGACAGATGATGAGCAAGATGCTGCAGAA 1910
351 GGAATGCAACGACCTTCAGGACAGATGATGATGATGATGATGATGATGATGATGAT 410
1911 TCGTTGAAGTCAATCCTTGCA-----TCGGCGAAACGGCACATCGCTTTCATTGAA 1966
411 GGTGGATGAGAGCTTTTGTGAGAGTATGTGCTCTGATGAGCACATGCTTTATGAG 470
1967 CCATATCTTCTCTGACAGACTGACAAAGTCCAGGGGCTGACCTTAATACTGTGAG 2026

Db 471 CCAAGGTCAAGTGTGCTCCAGSACCCAGTAAACATGTGGCCATGATACATTCTG 530
Qy 2027 GACACGCTGGATTCTT---CAACACCATTGAAGAGTCCCAACCAAGCCGTCACCCAG 2083
Db 531 GCCAATGACAGGTTCTTTTGAAGCGCAACCTCTCCAAAGTGCCACGATGGCCCTGACTGTA 590
Qy 2084 GGTTCGGGACATTTTGTCCCGCGGCAAAACATGTGTGTGGCACTGTGTGAAGAAAA 2143
Db 591 GGTGTGGGACATGTGTGAAGCTCCCAAGAGTATATCTATATCACTGAGCCCATTA 650
Qy 2144 GCCGACGCGATCCGCGGAACTGTGAAGGCCAGTACTGCTTGTGCGCCAGTTCATC 2203
Db 651 GCATTTGCTTTGTATTAAGCAATGAGAGAGTGTAAATCATATGTGACACAGTTTCTGCT 710
Qy 2204 CTGTAGATGACACACATGCCACATCATCTGTGATGAAGACCA 2248
Db 711 TTTCACACAGACCCACAGCACTGTGTGTCTGTGATGAGATGCA 755

RESULT 9 811 bp mRNA linear EST 18-AUG-2003
CF346432
LOCUS AGENCOURT 15227262 NICHD XGC SwbIn Xenopus tropicalis cDNA clone
DEFINITION IMAGE:6998096 5', mRNA sequence.

ACCESSION CF346432
VERSION CF346432.1 GI:33788300
SOURCE EST.
ORGANISM Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 811)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Rob Granger, University of Virginia
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov

Plate: LLM14686 Row: a Column: 07
High quality sequence start: 15
High quality sequence stop: 717.
Location/Qualifiers

FEATURES

1. 811
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6998096"
/issue_type="whole body"
/note="Vector: pEXpress-1. Site 1: EcoRV, Site 2: NotI;
Bulk tissue was collected from a whole 10 month old male
from the P6 strain. 1st strand cDNA was primed with a Not
I - oligo (dT) primer, double-stranded cDNA was cloned into
the Not I and EcoRV sites of pEXpress-1. Library was
size-selected for >1.5 kb fragments for an average insert
size of 1.92 kb. Library was normalized to Cots with a
180-fold reduction of actin. A non-normalized version of
this library is also available (NICHD XGC SwbIn). Library
was constructed by Open Biosystems (Huntsville, AL).
PLEASE NOTE: This library contains high level of
contamination by worm"

ORIGIN

Query Match 1.7%; Score 104.4; DB 7; Length 811;
Best Local Similarity 50.5%; Pred. No. 2,2e-18;
Matches 309; Conservative 0; Mismatches 296; Indels 7; Gaps 2;

Qy 1611 CCCTTGCCCAACAGAGGTGGAACCTTGGGGCTTGCACAGAGATCTCACCACAGTATACC 1670
Db 149 CCACCGCTGATTAATACCTTTACTTTAGGGCTCCCAACAGAGACCCACACTAGATGC 208
Qy 1671 TACCAAGACCTCATTCGATGATGAAGCTGGGGAAGTGTCACTTCAAGAACTGCAAGCA 1730
Db 209 TACAAAAAATAATAGATACCATATAAAACGAGATCTCTCTTCAAGATGTGAATAACA 268
Qy 1731 TTCTTTGTGATGAATATAGCTGGGACTTAACCCGTGAGATGAAAAACAGCTACTTAAC 1790
Db 269 TTCAATATGATGATGATGATGTGGCTCTCCAGGAGACCTGAGGCTATCATTCCTTC 328
Qy 1791 ATTGCAAGAGTTCATCTGACCACTGACATGCTGTGATGAAGAGTCTACAGCCAGAT 1850
Db 329 ATGTGAAACAATCTTTTAAGCAATGACATGACATGAGGCTGAGAAATGCCACATCTTGAT 388
Qy 1851 GGTGCAAAACCTGATCATACGAAGCAGCTGAGAGTATGAGCAAAAGATCGTCGAGAA 1910
Db 389 GGAATGCGACAGACCTTCAGGAGAGGTGTGACCTGTTGAAGAAAAAGATTGGGCTGCT 448
Qy 1911 TCCGTTGAAGTTCAATCTTGGCA---TCGGCGGAAACGGCACTGCTTCACTTGA 1966
Db 449 GGTGGATTAAGCTGTTTGTGTGAGGTATGTGCTGATGGGCACATACCTTTTAATGAG 508
Qy 1967 CCATCATCTTCTCTGTCAGAGCTGACCAAGTTCAGAGCCCTGACCCCTAAACTGTGAG 2026
Db 509 CCAGGTCAAGTCTGTGCTTCAGAGACCCGAGTAAACATTTGCCATGTATACATTTCTG 568
Qy 2027 GACACGCTGATCTT---CAACACCATGCAAGAGTCCCAACCCAGCGCTGACCCAG 2083
Db 569 GCCAATGACAGGTTCTTTTGAAGCGCAACCTCTCCAAAGTCCCAAGATGAGCCCTGACTGTA 628
Qy 2084 GGTTCGGGACATTTGTCCCGCGGCAAAACATGTGTGTGTGGCACTGTGTGAAGAAAA 2143
Db 629 GGTGTGGGACATGTGATGAGCTCCCAAGAGTATATCTATATCACTGAGCCCATTA 668
Qy 2144 GCCGACGCGATCCGCGGAACTGTGAAGGCCAGTACTGCTTCTGCGCCAGGTTCATC 2203
Db 689 GCATTTGCTTTGTATTAAGCAATGAGAAAGTGTAAATCATATGTGACAGTTTCTGCT 748
Qy 2204 CTGTAGATGCAC 2215
Db 749 TTTCACACAGCAC 760

RESULT 10 896 bp mRNA linear EST 26-SEP-2003
CF592490
LOCUS AGENCOURT 15680772 NICHD XGC SwbIn Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7024229 5', mRNA sequence.

ACCESSION CF592490
VERSION CF592490.1 GI:36345132
SOURCE EST.
ORGANISM Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-remail.nih.gov
Tissue Procurement: Rob Granger, University of Virginia

CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM14753 row: f column: 03
 High quality sequence stop: 708.
 Location/Qualifiers

FEATURES

1..896
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7024229"
 /cissue_type="whole body"
 /clone_lib="NICHID XGC SbdIN"
 /note="Vector: pEXpress-1; Site 1: EcoRV; Site 2: NotI;
 Bulk tissue was collected from a whole 10 month old male
 from the F6 strain. 1st strand cDNA was primed with a Not
 I - oligo(dT) primer, double-stranded cDNA was cloned into
 the Not I and EcoRV sites of pEXpress-1. Library was
 size-selected for >1.5 kb fragments for an average insert
 size of 1.92 kb. Library was normalized to Cots with a
 180-fold reduction of actin. A non-normalized version of
 this library is also available (NICHID XGC SbdI). Library
 was constructed by Open Biosystems (Huntsville, AL).
 PLEASE NOTE: This library contains high level of
 contamination by worm"

ORIGIN

Query Match 1.7%; Score 104.4; DB 7; Length 896;
 Best Local Similarity 50.5%; Pred. No. 2,3e-18;
 Matches 309; Conservative 0; Mismatches 296; Indels 7; Gaps 2;

1611 CCCTTCGCAAGAGGTGAACTTGGGGCTTGGCAAGATCTCAACCACTGAATACC 1670
 143 CCACCGCTGATTAATCTTATCTTAGGGCTCCCAAGAGCAAGCCCACTAGATGC 202
 1671 TACCAAGAGCTCAATGCGATGATGAGCTGGGGAAGTCTCAATCAAGACTGCAAGGA 1730
 203 TACAAAAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 262
 1731 TCTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1790
 263 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 322
 1791 ATTGCAAGAGTCACTGACCAATGACATGATGATGATGATGATGATGATGATGATGATGAT 1850
 323 ATGTGGAACAATCTTCTTAAGCAATGACATGACATGACATGACATGACATGACATGACATGACAT 382
 1851 GGTGCAAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1910
 383 GGAAATGCAAGACCTTCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
 1911 TCGGTTGAAGTCAAAATCTTGGCA---TCGGCGGAAACGGACATCGCTTCAATGTA 1966
 443 GGTGGGATGAGCTGTTGTTGAGGATATGTTGCTTGAAGGCACTAGCTTTTAATGAG 502
 1967 CCATCATCTTCTCTGTCAGAGTGAACAAAGTCCAGGCTGACCCCTTAAACTGTGAG 2026
 503 CCAGGATCAAGTCTGATCTTCAAGACCGAGTAAACATTTGGCCATGATACATTTCTG 562
 2027 GACCAAGCTGATCTT---CAACACATGCAAGAGTCCCAACCCAGCCGTCAACCG 2083
 563 GCCAATGACAGGTTCTTTGACGCAACCTTCCAAAGTCCCAACGATGCGCTGATGTA 622
 2084 GGTGGGAGCTTTGTCGCGCGCAAAACATGTTGTTGGGCAACTGTGAAGGAAA 2143
 623 GGTGGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
 2144 GCCGAGCCATCCGCGGAACGTGGAAGGCCAGTGAATGCTCTTGGCCAGGTTTCATC 2203
 683 GCATTTGCTTTGATGAAGCAATGAGAGGTGATTAATCATATGAGCAAGTTTCTGCT 742

QY 2204 CTGTAGATGCAC 2215
 Db 743 TTTTCACAGCAC 754

RESULT 11

CR696680 1177 bp mRNA linear HTC 19-AUG-2004
 LOCUS Tetraodon nigroviridis full-length cDNA.
 DEFINITION
 ACCESSION CR696680
 VERSION CR696680.1 GI:51194589
 KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE

1 (bases 1 to 1177)
 Genoscope.
 Direct Submission
 Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
 2 rue Gaston Creneau, CP 5706 - 91057 EVRY cedex - FRANCE
 (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 The sequences are based on single pass reads.
 More information available at
<http://www.genoscope.cns.fr/tetraodon>.

COMMENT

FEATURES

source

1..1177
 /organism="Tetraodon nigroviridis"
 /mol_type="mRNA"
 /db_xref="taxon:9983"
 /cissue_type="fyes"

ORIGIN

Query Match 1.7%; Score 104.4; DB 3; Length 1177;
 Best Local Similarity 49.7%; Pred. No. 2,5e-18;
 Matches 324; Conservative 0; Mismatches 321; Indels 7; Gaps 2;

1603 TAATGCAACCTTTCGCAACAAAGGTGAACTTGGGGCTTGGCAAGATCTTCAACAC 1662
 336 TATTCATATCCGGCCGCGACCGGTATTTTACCTGGGGCTGCCCCAAGAGCACTCTTC 395
 1663 TGAATCACTTACCAAGAGCTCAATGCGATGATGATGATGATGATGATGATGATGATGATGAT 1722
 396 TGGGTTGCTACCAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 455
 1723 GCAAGGCAATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1782
 456 TGAAGACCTTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 515
 1783 TTAACCAATTCGCAAGAGTCACTGACCAATGATGATGATGATGATGATGATGATGATGATGAT 1842
 516 ACTCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 575
 1843 GCCCAAGTGTGCAAAACCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1902
 576 TTTTGAATGCAACCGCCCGACCTGCAAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 635
 1903 CTGAGAAATCCGTTGAAGTCAAAATCTTGGCA---TCGGCGGAAACGGACATGCTGCT 1958
 636 TGGAGGCTGTGAGATCCAGCTGTTTGTGCGAAGATGATGATGATGATGATGATGATGATGATGAT 695
 1959 TCATTGAACATCATCTTCTCTGTCAGAGTGAACAAAGTCCAGGCTGACCTTAATA 2018
 696 TCAAGAGCCCGGTTCCAGTTTGTGTCAGAACCGGATGAAAGACCTTGGCCAAAGACA 755
 2019 CTGTGAGAGCAACGCTGATCTTCAAC---ACATGGAAGAGTCCCAACCCAGCGCG 2075
 756 CCATATTTGCCAACCGGCTTCTTGAAGGAGCTTCCAAATGTCACCAACCGGCC 815
 2076 TCACCAAGGTTTGGGCACTTGTCCCGCGCAAAACATGATGATGATGATGATGATGATGATGAT 2135

Db 816 TGACTGTGGCGGTGGGCACTGTATGAGAGCCAAAGAGTCATGATTTCTATACGGGCG 875
Qy 2136 AAGAAAAGCCGACGCCATCCGGGAACTGTGAAGGCCCATGACTGTTCTTGCCCA 2195
Db 876 CACACAAGGCTCTGGCTTTGGCCAAAGCTATCGAGGAGGGCGTGAATCATCATGTGACCG 935
Qy 2196 GTTCCATCCTGTAGTGCACACATGCGACCATCTGTTGATGTAAGCAGC 2247
Db 936 TCTGTGCTTCCAGACAGCCACAGTCATCTTGTGTGCGAGGAGCG 987

RESULT 12
CR690088 1178 bp mRNA linear HTC 19-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION CR690088
VERSION CR690088.1 GI:51187995
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1178)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1..1178
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/cissue_type="Eyes"

ORIGIN
Query Match 1.7%; Score 104.4; DB 3; Length 1178;
Best Local Similarity 49.7%; Pred. No. 2.5e-18;
Matches 324; Conservative 0; Mismatches 321; Indels 7; Gaps 2;
Qy 1603 TAATGCACCCCTTGGCAACAGGCTGAACTTTGGGGCTTGGACAGATCCTCACAC 1662
Db 303 TATTCAATCCCGGCGCGGACCGGTATTTACCTTGGGGCTGCCACAGAGACATCTCTC 362
Qy 1663 TGAATACCTTACCAAGAGCTCATTCGATGATGAAAGCTGGGGAAGTGTCACTCAAGACT 1722
Db 363 TGGGTGTCAACAAGAGCTGATGATATTAACAAGATGAGAAGCTCGTTCACTACG 422
Qy 1723 GCAAGGATTTCTTTGGATGATAGTGGGACTAACCGGTGACATGAAAAAGTACT 1782
Db 423 TGAAGACTTCAACATGAGCAATATGTGTCTCCCGAGAGCACCCCGAGAGTACC 482
Qy 1783 TTAACAATTCGCAAGAAGTTCATGACCATGACATCGTTGATGAAGAGTCTACA 1842
Db 483 ACTCCCTTATGTGAGCAACTTCTTCAAGACATGACATCAAGCGGAAAAACATCACA 542
Qy 1843 GCCCAGATGTGCAAACTCTGATCATACAGACAGCTGCAAGTATGAGCAAAAGATCG 1902
Db 543 TTTTAGACGGCAACGCGCGGACCTGACAGAGAGTGCAGAGCTTCGAGAGAAATGG 602
Qy 1903 CTGCAGATCCGTTGAAGTTCAATCCTTGGCA----TGGCGGAAACGGCAATGCTT 1958
Db 603 TGGAGGCTGTGGATTCACAGCTGTTGTGTGCGAGGAGTTGGAACGGACGACATTTGCC 662
Qy 1959 TCATTGAACCATCATCTTCTGTCTGTCAGACTGACAAAGTCCAGGCGCTGCACCTTAAA 2018
Db 663 TTAAGAGCCGCGTTCAGTTGTGTGTCAGAAACGGGGGTGAAGACCTTGGCCAGAGCA 722

Qy 2019 CTGTGAGAGCAACGCTCGATTCTTCAAC---ACCATGAAAGAGTCCCAACCCAGCG 2075
Db 723 CCATCATTTCCCAACGACGCTTTGTGAGCGGAGACTTCCAAAGGCCCAACGCCCC 782
Qy 2076 TCACCCAGAGTTTGGGCACTTTGTCCCGCGCAAAACATCGTTGTGTGCAACTGTG 2135
Db 783 TGACTGTGGCGGTGGGCACTGTTCATGAGAGCCCAAGAGTCAATGATTTCTATACGGCG 842
Qy 2136 AAGAAAAGCCGACGCGCATCCGGGAACTGTGGAAGGCCCAAGTACTGCTTCTTGCCAG 2195
Db 843 CACACAAGGCTCTGGCTTTGGCCAAAGCTATCGAGAGGCGCTTATCATCATGTGACCG 902
Qy 2196 GTTCCATCCTGTAGTGCACACATGCGACCATCTGTTGATGTAAGCAGC 2247
Db 903 TTTGTGCTTCCAGACAGCCACAGTCATCTTGTGTGCGAGGAGCG 954

RESULT 13
CR701935 1206 bp mRNA linear HTC 19-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION CR701935
VERSION CR701935.1 GI:51199844
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1206)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1..1206
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/cissue_type="Eyes"

ORIGIN
Query Match 1.7%; Score 104.4; DB 3; Length 1206;
Best Local Similarity 49.7%; Pred. No. 2.6e-18;
Matches 324; Conservative 0; Mismatches 321; Indels 7; Gaps 2;
Qy 1603 TAATGCACCCCTTGGCAACAGGCTGAACTTTGGGGCTTGGACAGATCCTCACAC 1662
Db 348 TATTCAATCCCGGCGCGGACCGGTATTTACCTTGGGGCTGCCACAGAGACATCTCTC 407
Qy 1663 TGAATACCTTACCAAGAGCTCATTCGATGATGAAAGCTGGGGAAGTGTCACTCAAGACT 1722
Db 408 TGGGTGTCAACAAGAGCTGATGATATTAACAAGATGAGAAGTCTGTTCCGTACG 467
Qy 1723 GCAAGGATTTCTTTGGATGATAGTGGGACTAACCGGTGACATGAAAAAGTACT 1782
Db 468 TGAAGACTTCAACATGAGCAATGCTGTGTCTCCCGAGAGCACCCCGAGAGTACC 527
Qy 1783 TTAACAATTCGCAAGAAGTTCATGACCATGACATCGTTGATGAAGAGTCTACA 1842
Db 528 ACTCCCTTATGTGAGCAACTTCTTCAAGACATGACATCAAGCGGAAAAACATCACA 587
Qy 1843 GCCCAGATGTGCAAACTCTGATCATACAGAGAGCTGCAAGTATGAGCAAAAGATCG 1902
Db 588 TTTTAGATGGCAACCGCGCGGACCTGACAGAGAGTGCAGAGCTTCGAGAGAAATGG 647
Qy 1903 CTGCAGATCCGTTGAAGTTCAATCCTTGGCA----TGGCGGAAACGGCAATGCTT 1958

Db 648 TGGAGCTGGTGGATCCAGCTGTTGTGGAGGAATTGGACCGGACCGCCATTGGCT 707
 Qy 1959 TCATTGAACCATCATCTTCTCTGTGCAGACTGACAAAGGTCCAGGGCTGCACCTTAATA 2018
 Db 708 TCAGAGAGCCGGTTCCACTTTGTGTCCAGAACCGGGGTGAAGACCTCGGCCAAGAGACA 767
 Qy 2019 CTGTGGAGACAAACGCTCGATTCTTCAAC--ACCATGGAAGAGGTCCCAACCCAGCGCC 2075
 Db 768 CCATCATTTGCCCAACGGCGCCCTCTTTGACGGGAGACCTCTCCAAAGTCCCAACAGCGCC 827
 Qy 2076 TCACCCAGGGTTTGGGCACTTTGTCCCGGCGCAAAACATGCTTGTGGCAACTGTG 2135
 Db 828 TCATCTGTGGCGCTGGGCACTGTATGACACGCCAAGAGGTCAATGATTCTATCAGGAGCG 887
 Qy 2136 AAGAAAAGCCGACCGCATCCGCGGAACCTGTGAAAGCCCACTGACTGCTTTGCCAG 2195
 Db 888 CACACAGGCTCTGGGCTTTGGCCAAAGCTATCGAAGAGGCGGTGAATCATGTGAGACG 947
 Qy 2196 GTTCCATCTGTAGATGCACACATGCCACCATCTGTTGATGAGACG 2247
 Db 948 TCTCTGCTTCAGACGACCAACGATCATCTCTGTGTGCGACGAGAGCG 999

RESULT 14
 LOCUS BM017859 818 bp mRNA linear EST 30-OCT-2001
 DEFINITION 603645335F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5427056 5',
 mRNA sequence.
 ACCESSION BM017859
 VERSION BM017859.1 GI:16532213
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1892 row: e column: 09
 High quality sequence stop: 803.
 Location/Qualifiers

FEATURES
 source 1..818
 Location/Qualifiers

1..818
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5427056"
 /tissue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH_MGC_98"
 /note="Organ: Brain; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

Query Match 1.7%; Score 103.8; DB 4; Length 818;
 Best Local Similarity 49.5%; Pred. No. 3.3e-18;
 Matches 327; Conservative 0; Mismatches 327; Indels 7; Gaps 2;

Qy 1610 ACCCTTCCGCAACAAAGGTGAACTTTGGGGCTTGCAACAGATCTCTACCACTGACGTAAC 1669
 Db 78 ACCCGAGGCGAGAGAAAGTACTTTCACCTCTGGGGCTCCCACTGGAGTACCCCACTTGGCTG 137
 Qy 1670 CTACCAAGAGCTCATTCGATGATGATGAAGCTGGGGGAAGTGCATTCAGAACTGCAGAGC 1729
 Db 138 CTACAGAAAGCTGATTTGAATCTATTAAGATGGGAGCTGTCTTTAAATATGTGAAGAC 197
 Qy 1730 ATTCTTGTGATGAATAAGTGGGACTAACCCTGACGATGAATAACAGCTACTTTAAAC 1789
 Db 198 CTTGAACATGATGATGATGATGAGGCTTCTCTGAGACACCGGAGATTAACACTCTCTT 257
 Qy 1790 CATTGGCAAGAAGTTCATGACACATGACATGATGATGATGAAGGCTTACAGCCCA 1849
 Db 258 CATGTGAAACAACCTTCTTCAAGCAATTTGACATCCACGAGAAACACCCACATTTGGA 317
 Qy 1850 TGGTCAAAACCCGTATTCATACGAAGCAGCTGACAGATGATGAGGCAAGATGCTGCAGA 1909
 Db 318 TGGGAATGACATGACCTACAGGAGAAATGTGATGCTTTGAAGAAAGATCAAGGCTGC 377
 Qy 1910 ATCCGTTGAAGTTCAAAATCTTGGCATGCGCGG---AAACGGCAATGCTTTGATTGA 1965
 Db 378 AGGTGGATTCAGAGTATTGTTGGAGGATCGGCGCTGATGAGACATTTGCCCTTCAAGA 437
 Qy 1966 ACCATCATCTTCTCTGTGAGACTGACAAAGTCCAGGCGCTGCACCTTAATACTGTGGA 2025
 Db 438 GCCAGGCTTCAGTCTGTGTGTCGAGACCCGTGTGAAGCGCTGGCATGATATCCATCTT 497
 Qy 2026 GGACAAACGCTCGATTCTTCAA---CACCATGAAGAAGTCCCAACCCAGCGCTGACCCA 2082
 Db 498 GGCCAAATGCTAGTGTCTTTCATGAGAACTACCAAGGTGCCACCATGAGCTTGAACGCT 557
 Qy 2083 GGGTTGGGCACTTGTGCTCCCGCGCAAAACATGCTGTTGTGTGCACTGTGTAAGAA 2142
 Db 558 GGGGGTGGGCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
 Qy 2143 AGCCGACCCCATCCCGGGAACCTGTGAGAGCGCCAGTGCCTGTTTCCCAAGTTTCAT 2202
 Db 618 GGCATTTCTCTGTACAAAGCCCATGAGAGGAGAGTGAACCAATGTGACCGTGTCTGC 677
 Qy 2203 CCTGTAGATGACAAACATGACCATCATCTGTTGATGAAGAGCAGATTCACAGCTGGA 2262
 Db 678 CTTCCAGACAGCATCCCGGACCGGGTTGTGTGTGAGAGAGATGCCACTTGGAGACTGA 737
 Qy 2263 A 2263
 Db 738 A 738

RESULT 15
 LOCUS AL656223 662 bp mRNA linear EST 07-NOV-2003
 DEFINITION AL656223 XGC-neurula Xenopus tropicalis cDNA clone tneu03507 5',
 mRNA sequence.
 ACCESSION AL656223
 VERSION AL656223.2 GI:38223958
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 REFERENCE 1 (bases 1 to 662)
 AUTHORS Crotting,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
 JOURNAL Unpublished (2003)
 COMMENT On Dec 13, 2001 this sequence version replaced gi:1768464.
 Contact: Huckle E
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu035007.plksp6
Sequencing primer: SP6.
Location/Qualifiers

FEATURES

source

1..662

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TNeu035007"

/dev_stage="neurula"

/lab_host="Escherichia coli DH10B"

/clone_lib="XGC-neurula"

/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA

was oligo dt primed from 5ug of poly A+ RNA from neurula.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

```
Query Match      1 7%; Score 103.2; DB 1; Length 662;
Best Local Similarity 52.0%; Pred. No. 4.5e-18;
Matches 282; Conservative 0; Mismatches 253; Indels 7; Gaps 2;

QY 1611 CCCTTCGCCAAGAGGAGGAGAACTTGGGGCTTGCACAGAGATCCTCACCAGTAGTACC 1670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 111 CCAACCGCTGATTAATTAATCTTCTTAGGGCTCCCAAGAGAGACCCCACTAGGATGC 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1671 TACCAAGAGCTCATTCGATGATGAGAGTGGGAAAGTGTCAITCAAGAACTGCAAGCA 1730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 TACAATAAACTAATAGATACCATATAAAAGGAGATCTCTCTCAAGTATGTGAAAACA 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1731 TTCTGTGGATGATGATAGCTGGAGACTAACCCGAGAGATGAAAACGCTACTTTAAACC 1790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 231 TTCAATATGATGATGATGATGAGGCTTCCAGGAGACACCTGAGAGCTATCATTCCTTC 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1791 ATTCCAAAGAGTTCACATGACCAATGCAATGATGTTGATGAAGAGTCTACAGCCAGAT 1850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 291 ATGTGGAACAATCTCTTAAGCAATGCAATGACGCTGAGATGCCACATCTTGGAT 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1851 GGTGCAAACTGATTCATATGCAAGCAGTGCAGAGTATGAGCAAAAGATCGCTGCAGAA 1910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 351 GGAAATGCCACAGACCTTCAGGACAGTGTGACTGTGTAAGAAAAGATCGGGCTGCT 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1911 TCCGTTGAAGTTCGAATCCTTGG----CATGGCGGAAACGGACATCGCTTTCATTGAA 1966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 GGTGGATTGAGCTGTTGTGAGGTATGTTCTGATGGCCACATAGCTTTTAATGAG 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1967 CCATCATCTTCTCTGTCAGAGACTGACAAAGTTCAGAGCGCTGCACCCCTAAACTGTGAG 2026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 CCAAGGTCAGTCTGTGTCACGAGACCCAGTAAACATGTGCCATGATACATCTTCTG 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2027 GACAAAGCTGATTTCTT---CAACACCATGCAAGAGGTCCCAACCCAGCCGTACCCAG 2083
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 531 GCCAATGCAAGGTTCTTGAAGGCAACCTCCAAAGTGCCAAAGATGAGCCCTGACTGTA 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2084 GGTTCGGGCACTTTGTCGGCGCGCAAAACATGCTGTGTGTGGCAATGTTGAAGGAAA 2143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 591 GGTGTGGGACTGTGTGATGATCCCAAGAGTCAATGATTCTAATCACTGAGGCCATAAA 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2144 GC 2145
    ||
DB 651 GC 652
```

Search completed: March 10, 2005, 08:24:38
Job time : 16585 secs

This Page Blank (uspto)

CC amino-acid and nucleic acid productivity. (Updated on 11-SEP-2003 to
CC standardise OS field)

XX Sequence 661 AA;

Query Match 100.0%; Score 3342; DB 4; Length 661;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDHKLQRIIRDIGEDENI VAAACATRLRLVLTQDKVDROSLDDPDLKGTETGGM 60
DB 1 MDHKLQRIIRDIGEDENI VAAACATRLRLVLTQDKVDROSLDDPDLKGTETGGM 60
QY 61 FOIIVPGDVHVFKELDATSKDI AVSTEQDKVANNANMFSAVKVLADIFVPLPI 120
DB 61 FOIIVPGDVHVFKELDATSKDI AVSTEQDKVANNANMFSAVKVLADIFVPLPI 120
QY 121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180
DB 121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180
QY 181 REGNEFLGAGIGMANVFPPTLVNGYDVAATMTAGEMPWMSLFGLDVAAGYQGVLPVLV 240
DB 181 REGNEFLGAGIGMANVFPPTLVNGYDVAATMTAGEMPWMSLFGLDVAAGYQGVLPVLV 240
QY 241 VSMILATIEKFLHKRLMGTAFLITPVLTLLTGFTITFAIGPMRWGDLIAHGLQGLY 300
DB 241 VSMILATIEKFLHKRLMGTAFLITPVLTLLTGFTITFAIGPMRWGDLIAHGLQGLY 300
QY 301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANINOGAACLAIV 360
DB 301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANINOGAACLAIV 360
QY 361 FFLAKSEKIKGLAGASGSAVAGITEPAIFGVNRLRMPFYIGITAAIGALIALFDIK 420
DB 361 FFLAKSEKIKGLAGASGSAVAGITEPAIFGVNRLRMPFYIGITAAIGALIALFDIK 420
QY 421 AVAALGAGFLGVSDIDAPDMNMFVCAVVTPIAAGAAIAYGLVYRRNGSIDDPATAAP 480
DB 421 AVAALGAGFLGVSDIDAPDMNMFVCAVVTPIAAGAAIAYGLVYRRNGSIDDPATAAP 480
QY 481 VPAGTTKAEAPAEESNDSTIIQAPLTGEAIALSSVSAMPASGKLSGSAIVPTKGL 540
DB 481 VPAGTTKAEAPAEESNDSTIIQAPLTGEAIALSSVSAMPASGKLSGSAIVPTKGL 540
QY 541 VSPVSGKIVVAPPSGHAFAVRTKAEDGSNDILMHIGFDTVNLNGTHFNPLKKQDEYVA 600
DB 541 VSPVSGKIVVAPPSGHAFAVRTKAEDGSNDILMHIGFDTVNLNGTHFNPLKKQDEYVA 600
QY 601 GELLCEPDIDAKAGYEVTTPPIVSNYKKTGPVNTYGLGEIEAGANLLNVAKKAVPAT 660
DB 601 GELLCEPDIDAKAGYEVTTPPIVSNYKKTGPVNTYGLGEIEAGANLLNVAKKAVPAT 660
QY 661 P 661
DB 661 P 661

```

RESULT 2
AAG92650
ID AAG92650 standard; protein; 661 AA.
XX
AC AAG92650;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6404.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis.
XX
OS Corynebacterium glutamicum.

XX

PN EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOM) KYOMA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senon A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.

XX N-PSDB; AAH67869.

PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.

PS Claim 17; SEQ ID NO 6404; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of Corynebacterium glutamicum, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from

CC Corynebacterium glutamicum, and identifying a homologue of a gene derived from

CC Corynebacterium glutamicum. Corynebacterium bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described in the

CC exemplification of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the European Patent Office

XX

SEQ Sequence 661 AA;

Query Match 99.0%; Score 3310; DB 4; Length 661;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 654; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MDHKLQRIIRDIGEDENI VAAACATRLRLVLTQDKVDROSLDDPDLKGTETGGM 60
DB 1 MDHKLQRIIRDIGEDENI VAAACATRLRLVLTQDKVDROSLDDPDLKGTETGGM 60
QY 61 FOIIVPGDVHVFKELDATSKDI AVSTEQDKVANNANMFSAVKVLADIFVPLPI 120
DB 61 FOIIVPGDVHVFKELDATSKDI AVSTEQDKVANNANMFSAVKVLADIFVPLPI 120
QY 121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180
DB 121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180
QY 181 REGNEFLGAGIGMANVFPPTLVNGYDVAATMTAGEMPWMSLFGLDVAAGYQGVLPVLV 240
DB 181 REGNEFLGAGIGMANVFPPTLVNGYDVAATMTAGEMPWMSLFGLDVAAGYQGVLPVLV 240
QY 241 VSMILATIEKFLHKRLMGTAFLITPVLTLLTGFTITFAIGPMRWGDLIAHGLQGLY 300
DB 241 VSMILATIEKFLHKRLMGTAFLITPVLTLLTGFTITFAIGPMRWGDLIAHGLQGLY 300
QY 301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANINOGAACLAIV 360
DB 301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANINOGAACLAIV 360
QY 361 FFLAKSEKIKGLAGASGSAVAGITEPAIFGVNRLRMPFYIGITAAIGALIALFDIK 420
DB 361 FFLAKSEKIKGLAGASGSAVAGITEPAIFGVNRLRMPFYIGITAAIGALIALFDIK 420
QY 421 AVAALGAGFLGVSDIDAPDMNMFVCAVVTPIAAGAAIAYGLVYRRNGSIDDPATAAP 480

```

Db 421 AVALGAAGFLGVSIADPDMWFLVCAVVTFFIAFGAATAYGLYVRNGSIDPDATAP 480
Qy 481 VPAGTTKAEADAPAEFSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSCVAIVPTKGQI 540
Db 481 VPAGTTKAEADAPAEFSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSCVAIVPTKGQI 540
Qy 541 VSPVSGKTIIVAFPSGIAFAVRTKAEDEGSNDILMHIGFPTVNLNGTHFNPCLKQGBEYVA 600
Db 541 VSPVSGKTIIVAFPSGIAFAVRTKAEDEGSNDILMHIGFPTVNLNGTHFNPCLKQGBEYVA 600
Qy 601 GELLCFDPIDAIKAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNVAKKEAVPAT 660
Db 601 GELLCFDPIDAIKAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNVAKKEAVPAT 660
Qy 661 P 661
Db 661 P 661

RESULT 3
AAB6707
ID AAB6707 standard; protein; 468 AA.
XX
AC AAB6707;
XX
DT 09-APR-2001 (first entry)
XX
DE C.glutamincum phosphoenolpyruvate protein #1.
XX
KM Phosphoenolpyruvate; sugar phosphotransferase system; PTS.
XX
OS Corynebacterium glutamicum.
XX
XX WO200102583-A2.
XX
PD 11-JAN-2001.
XX
PF 27-JUN-2000; 2000MO-IB000973.
XX
PR 01-JUL-1999; 99US-0142691P.
XX
PR 23-AUG-1999; 99US-0150310P.
XX
PR 03-SEP-1999; 99DE-01042095.
XX
PR 03-SEP-1999; 99DE-01042097.
XX
PA (BAD1) BASF AG.
XX
PI Pompejus M, Kroegeer B, Schroeder H, Zeider O, Haberhauer G;
XX
DR WPI; 2001-080989/09.
XX
XX Corynebacteiium glutamicum nucleic acids encoding phosphoenolpyruvate;
XX
PT sugar phosphotransferase system proteins or their portions, useful for
XX
PT typing or identifying C. glutamicum or related bacteria, and as markers
XX
PT for transformation.
XX
PS Claim 4; Page 101-102; 144pp; English.
XX
XX The present invention relates to Corynebacteiium glutamicum
XX
CC phosphoenolpyruvate; sugar phosphotransferase system (PTS) proteins. The
XX
CC PTS nucleic acids and proteins are useful in the identification of
XX
CC microorganisms which can be used to produce fine chemicals, for
XX
CC modulating fine chemical production in C. glutamicum or related bacteria,
XX
CC the typing or identification of C. glutamicum or related bacteria, as
XX
CC reference points for mapping C. glutamicum genome, and as markers for
XX
CC transformation
XX
SQ Sequence 468 AA;
XX
Query Match 69.8%; Score 2332; DB 4; Length 468;
Best Local Similarity 98.5%; Pred. No. 2,1e-222;
Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 194 MANVPTLVNGIVDAATMTATGEMPMWSLFGLDVAQAGYQGTVPVLVWSHILATITEKFLH 253

Db 1 MANVPSLVNGIVDAATMTATGEMPMWSLFGLDVAQAGYQGTVPVLVWSHILATITEKFLH 60
Qy 254 KRLMGTADEPLITPVLTLLTGTFLTPTAIGPARMWGDLAHGLOGLYDFGPGVGLLFGI 313
Db 61 KRLMGTADEPLITPVLTLLTGTFLTPTAIGPARMWGDLAHGLOGLYDFGPGVGLLFGI 313
Qy 314 VSPPIVITGLHOSFPPIBELFNQGSPTPATASMANIQAACLAVFPLASEKLGIA 373
Db 121 VSPPIVITGLHOSFPPIBELFNQGSPTPATASMANIQAACLAVFPLASEKLGIA 373
Qy 374 GASGSVAIVGITTEPAIFGVNLRMPFYIGTAAIGALLILFDIKAVLAGAGFLGYV 433
Db 181 GASGSVAIVGITTEPAIFGVNLRMPFYIGTAAIGALLILFDIKAVLAGAGFLGYV 433
Qy 434 SIDAPDMWFLVCAVVTFFIAFGAATAYGLYVRNGSIDPDATAPVPAGTTKAEADAP 493
Db 241 SIDAPDMWFLVCAVVTFFIAFGAATAYGLYVRNGSIDPDATAPVPAGTTKAEADAP 493
Qy 494 AEFSDNSTIIQAPLTGEAIALSSVSDAMFASGKLGSCVAIVPTKGQI VSPVSGKTIIVAF 553
Db 301 AEFSDNSTIIQAPLTGEAIALSSVSDAMFASGKLGSCVAIVPTKGQI VSPVSGKTIIVAF 553
Qy 554 SGHAFAVRTKAEDEGSNDILMHIGFPTVNLNGTHFNPCLKQGBEYVAGELLCFDPIDAIK 613
Db 361 SGHAFAVRTKAEDEGSNDILMHIGFPTVNLNGTHFNPCLKQGBEYVAGELLCFDPIDAIK 613
Qy 614 AAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNVAKKEAVPATP 661
Db 421 AAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNVAKKEAVPATP 661

RESULT 4
ABG80325
ID ABG80325 standard; protein; 468 AA.
XX
AC ABG80325;
XX
DT 15-NOV-2002 (first entry)
XX
XX C. glutamicum metabolic pathway (MP) protein #5.
XX
DE C. glutamicum metabolic pathway (MP) protein #5.
XX
KM Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;
XX
KM cofactor; nucleotide; nucleoside; creatinase; fine chemical production;
XX
KM organic acid; non-proteinogenic amino acid; purine base; carbohydrate;
XX
KM pyrimidine base; lipid; unsaturated fatty acid; diol; polyketide;
XX
KM aromatic compound; food industry; animal feed; cosmetic industry;
XX
KM pharmaceutical industry; enzyme.
XX
OS Corynebacterium glutamicum ATCC 13032.
XX
XX WO200251231-A1.
XX
PN WO200251231-A1.
XX
PD 04-JUL-2002.
XX
PF 22-DEC-2000; 2000MO-EP013143.
XX
XX 22-DEC-2000; 2000MO-EP013143.
XX
PR 22-DEC-2000; 2000MO-EP013143.
XX
PA (BAD1) BASF AG.
XX
PI Pompejus M, Kroegeer B, Zeider O, Schroeder H;
XX
DR WPI; 2002-643289/59.
XX
DR N-PSDB; ABG65346.
XX
XX New metabolic pathway genes of Corynebacterium glutamicum for producing
XX
PT fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,
XX
PT cofactors or enzymes used in food, feed, cosmetics or pharmaceutical
XX
PT industries.
XX
PS Claim 18; Page 111-113; 176pp; English.
XX

CC The present invention relates to the isolation of Corynebacterium
 CC glutamicum metabolic pathway (MP) proteins, and the polynucleotide
 CC sequences encoding them. The MP proteins are enzymes involved in the
 CC metabolism of molecules important for the normal functioning of cells
 CC (e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or
 CC trehalose). The polynucleotide sequences encoding the MP proteins are
 CC useful for producing fine chemicals, particularly organic acids, non-
 CC proteinogenic amino acids, purine and pyrimidine bases, nucleosides,
 CC nucleotides, lipids, (un)saturated fatty acids, diols, carbohydrates,
 CC aromatic compounds, vitamins, cofactors, polyketides and enzymes. The
 CC fine chemicals are useful in the food, animal feed, cosmetic or
 CC pharmaceutical industries. ABG80321-ABG80343 represent the C. glutamicum
 CC MP proteins of the invention
 XX
 SQ Sequence 468 AA;

Query Match 69.8%; Score 2332; DB 5; Length 468;
 Best Local Similarity 98.5%; Pred. No. 2.1e-222;
 Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 194 MAMVPEPTLVNGYDVAATMAGMPWMSLFGLDVAQAQGYGTLPVLVVSMTATTEKPLH 253
 DB 1 MAMVPEPTLVNGYDVAATMAGMPWMSLFGLDVAQAQGYGTLPVLVVSMTATTEKPLH 60
 QY 254 KRLMGADPLTPVLTLLTGTFTFIAIGPAMRWGDLAHGLQGLYDPGGPVGGLFGL 313
 DB 61 KRLMGADPLTPVLTLLTGTFTFIAIGPAMRWGDLAHGLQGLYDPGGPVGGLFGL 120
 QY 314 VYSPVITGLHQSFPPIELFENOGGSFIFATASMANIAOGAACLAVFLAKSEKLGIA 373
 DB 121 VYSPVITGLHQSFPPIELFENOGGSFIFATASMANIAOGAACLAVFLAKSEKLGIA 180
 QY 374 GASGSVAIVGTEPAIFGNILRMPFYIGTAAIGALLLPDIKAVAGAFGLV 433
 DB 161 GASGSVAIVGTEPAIFGNILRMPFYIGTAAIGALLLPDIKAVAGAFGLV 240
 QY 434 SIDADPMVFLCAVTFEVIAGFAIAYGLYVRNRSIDPDATAAPVAGTTKAEAP 493
 DB 241 SIDADPMVFLCAVTFEVIAGFAIAYGLYVRNRSIDPDATAAPVAGTTKAEAP 300
 QY 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQOLVSPVSGKIVAFP 553
 DB 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQOLVSPVSGKIVAFP 360
 QY 554 SGHARAVRTKADGSDNDIMHIGFDTVNLNGTHFNPLKKOGDEYKAGELLCEPDIAIK 613
 DB 361 SGHARAVRTKADGSDNDIMHIGFDTVNLNGTHFNPLKKOGDEYKAGELLCEPDIAIK 420
 QY 614 AAGYEVTTPIVVSNYKKTGPVNTYIGLGEIAGANILNVAKKEAVPATP 661
 DB 421 AAGYEVTTPIVVSNYKKTGPVNTYIGLGEIAGANILNVAKKEAVPATP 468

RESULT 5
 AAB66708
 ID AAB66708 standard; protein; 362 AA.
 XX
 AC AAB66708;
 XX
 DT 09-APR-2001 (first entry)
 XX
 DE C-glutamicum phosphoenolpyruvate protein #2.
 XX
 KM Phosphoenolpyruvate; sugar phosphotransferase system; PTS.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN WO200102583-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-IB000973.
 XX

PR 01-JUL-1999; 99US-0142691P.
 PR 23-AUG-1999; 99US-0150310P.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042097.
 XX
 PA (BADI) BASF AG.
 PI
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Habenerauer G;
 DR WPI; 2001-080989/09.
 XX
 XX
 PT Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate;
 PT sugar phosphotransferase system proteins or their portions, useful for
 PT typing or identifying C. glutamicum or related bacteria, and as markers
 PT for transformation.
 XX
 PS Claim 6; Page 104-106; 144pp; English.
 CC
 CC The present invention relates to Corynebacterium glutamicum
 CC phosphoenolpyruvate; sugar phosphotransferase system (PTS) proteins. The
 CC PTS nucleic acids and proteins are useful in the identification of
 CC microorganisms which can be used to produce fine chemicals, for
 CC modulating fine chemical production in C. glutamicum or related bacteria,
 CC the typing or identification of C. glutamicum or related bacteria, as
 CC reference points for mapping C. glutamicum genome, and as markers for
 CC transformation
 CC
 SQ Sequence 362 AA;

Query Match 54.0%; Score 1806; DB 4; Length 362;
 Best Local Similarity 99.2%; Pred. No. 2.9e-170;
 Matches 359; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 300 YDFGPGVGGLLFGLVYSPVITVITGLHQSFPPIELFENOGGSFIFATASMANIAOGAACL 359
 DB 1 YDFGPGVGGLLFGLVYSPVITVITGLHQSFPPIELFENOGGSFIFATASMANIAOGAACL 60
 QY 360 VFPLAKSEKLGKLGAGASVSVALGTEPAIFGNILRMPFYIGTAAIGALLLPDI 419
 DB 61 VFPLAKSEKLGKLGAGASVSVALGTEPAIFGNILRMPFYIGTAAIGALLLPDI 120
 QY 420 KAVLAGAAGFLGVSIDADPMVFLCAVTFEVIAGFAIAYGLYVRNRSIDPDATPA 479
 DB 121 KAVLAGAAGFLGVSIDADPMVFLCAVTFEVIAGFAIAYGLYVRNRSIDPDATPA 180
 QY 480 PVPAGTTAKAEAPAEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQ 539
 DB 181 PVPAGTTAKAEAPAEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQ 240
 QY 540 LVPVSGKIVVAFPSGHAFAVTRKADGSDNDIMHIGFDTVNLNGTHFNPLKKOGDEVK 599
 DB 241 LVPVSGKIVVAFPSGHAFAVTRKADGSDNDIMHIGFDTVNLNGTHFNPLKKOGDEVK 300
 QY 600 AGEILLCEPDIDAIKAGYEVTTPIVVSNYKKTGPVNTYIGLGEIAGANILNVAKKEAVPA 659
 DB 301 AGEILLCEPDIDAIKAGYEVTTPIVVSNYKKTGPVNTYIGLGEIAGANILNVAKKEAVPA 360
 QY 660 TP 661
 DB 361 TP 362

RESULT 6
 ADK46609
 ID ADK46609 standard; protein; 627 AA.
 XX
 AC ADK46609;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Streptococcus pneumoniae protein, Seq ID No 3124.
 XX
 PF Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
 XX

```

XX OS Streptococcus pneumoniae.
XX PN US6699703-B1.
XX PD 02-MAR-2004.
XX PF 26-MAY-2000; 2000US-00583110.
XX PR 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX PR 30-JUN-1998; 98US-00107433.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewaeat CE;
XX DR WPI; 2004-212399/20.
XX DR N-PSDB; ADK43948.
XX PT New nucleic acid molecules and polypeptides useful for diagnosing,
XX PT preventing and treating pathological conditions resulting from bacterial
XX PT infection, e.g. Streptococcus pneumoniae infection, and in drug
XX PT screening.
XX PS Disclosure; SEQ ID NO 3124; 301bp; English.
XX CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
XX CC and polypeptides. The nucleic acids and proteins are useful for
XX CC diagnosing, preventing and treating pathological conditions resulting
XX CC from bacterial infection, such as S. pneumoniae infection. These may also
XX CC be used for drug screening procedures. The present sequence represents a
XX CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
XX CC data for this patent did not appear in the printed specification but was
XX CC obtained in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html.
XX SQ Sequence 627 AA;

Query Match      28.9%; Score 965; DB 8; Length 627;
Best Local Similarity 33.4%; Pred. No. 3.1e-86;
Matches 223; Conservative 141; Mismatches 245; Indels 58; Gaps 14;

QY 1 MHKDLAQRILRDIGEDNIVAAAHGATRLRLVKTQVDSQSLDDDDDKTFTGCM 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MNQELIAKKVIALGGRENNVSVAHCATRLRVWVQDEKINKKIVINLEKVKQCAFNSQ 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 FOIIVGPGVDVDFEKE--LDATSKDIATVSTEQLDKDVVANNANWFSRAVKVLADJFVPL 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 YQIIFGTGVNKNYDDEVVVLGLPTS-----SKDMLAEYAKQGNMFORAIRTFGDFVFI 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 IPIIVGGGLMANNVVAQDLFGPSLVEMFQISVAMINIMASAPFAFLPVLVGPT 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 116 IPIVIVATGIFMGVRGIFNLLEM-----PLPDPATVYQILIDTAFIILPGLVMS 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 ATKRFEGNEFLGAGIGMANVFPVLVNGYVYAATMTGEMPMMSLFGLDVAQAGYGTVP 237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 166 TRRFVEGNNPVGIVLGMMLVSGSLPNAWVA--OQGEVYANMFEPF-IPVVGILQGSVLP 221
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 VLVVSIILATIEKFLHKRLMGTAFLITPVLTLILGFLTFILIAIGAMVMVGDLHLHGQ 297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 222 APLIVGAKFEBAVKVVDVLDLVTTPVILLVNSIGLFGITGVFHVENVYIILATK 281
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 GYVDFGPGVGLLFGLVSPVITVTHAGSFPPIELFLNQGSGFIF-ATASMANIAGAA 356
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 282 AILSIPFGGFLIGVHQLIVSGVHHIFNLLEVOGLADHANPNFNAITTAAMTAQAA 341
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 357 CLAIFFLASSEKLGAGAGSVAVYGITRPAIFGVNLRKMPFYIGITGAIGALLIL 416
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 342 TVAVGVKTNPXKLTALFAPMAISAFILGITEPAIFGVNLFRRKPFSLTAGAIGGGLASI 401
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 417 FDIKAVLAGAGLGVISIDAPDMV-----MPLVCAVTFEVIAFGAIAVGLTVLR 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

DB 402 -----LGLAGTNGGNTIIPGMVLYVNGQLPQYLLMVAVSFALGFALTYMGY----- 449
QY 469 NSGIDPDATPAAPVPACTTAENE-APAEFSNDSTIIQAPLTGEALASSVSAMPASGL 527
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 450 --EDEVDAATAAKQAEVAEKEKEVAPALQNETLV--TPIVADVVALAVNDPVFSSGAM 505
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 528 GSGVAIVPTKQGLVSPGSKIVAVPSPGAFVARTKAEDGSNDILMHIGPDTVNLNGTH 587
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 506 GQGIIVAKSQQGVVYVAPADAEVSIAPFTGAFGLKTR---NGAEVLIHIGIDIVSNMG 561
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 588 FNPPLKQDEVYAGELCEFDIDAIAKAGYEVTTPIVSN---YKKTGPVNTYGLGEIRA 644
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 562 FEAKVYAQGNKVKRAGDVLGTFDSNKIATAAGLDJDTWIVYVNTADVASVAVAT---GSVAK 618
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 645 GANLNV 651
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 619 GDVAVIEV 625

RESULT 7
ADR94615
ID ADR94615 standard; protein; 643 AA.
AC ADR94615;
DT 16-DEC-2004 (first entry)
DE Novel S. pneumoniae protein sequence, SEQ ID 3250.
DE Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
KM bacterial infection.
XX OS Streptococcus pneumoniae.
XX PN US6800744-B1.
XX PD 05-OCT-2004.
XX PF 30-JUN-1998; 98US-00107433.
XX PR 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm LA, Bush D;
XX DR WPI; 2004-697205/68.
XX DR N-PSDB; ADR92012.
XX PT New isolated nucleic acid encoding a Streptococcus pneumoniae
XX PT polypeptide, useful for diagnosing, preventing and/or treating
XX PT pathological conditions resulting from the bacterial infection.
XX PS Disclosure; SEQ ID NO 3250; 151bp; English.
XX CC The invention relates to an isolated nucleic acid comprising a sequence
XX CC encoding a Streptococcus pneumoniae ADR9136polypeptide, or its
XX CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
XX CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642,
XX CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
XX CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or
XX CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
XX CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
XX CC hybridizable under high stringency conditions to the nucleotide sequence.
XX CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
XX CC Also included are a recombinant expression vector comprising the isolated
XX CC nucleic acid cited above operably linked to a transcription regulatory
XX CC element, a cell comprising the recombinant expression vector and a probe
XX CC comprising at least 20 consecutive nucleotides of the nucleotide
XX CC sequences as cited above. The methods and compositions of the present
XX CC invention are useful for the diagnosis, prevention and/or treatment of
XX CC pathological conditions resulting from bacterial infection by

```


DB 414 ARLIHVGDITVSLNKGFT--KVGAKKXKVGELGTFDSVITNSGDDTTMYIVN 471
QY 628 YKKTGPVNTYGLGEIEAGANLNV 651
DB 472 SKDYSEVITPTKIVTEGAALLTI 495

RESULT 9
AB847495
ID AB847495 standard; protein; 617 AA.
XX AB847495;
XX
XX 05-FEB-2002 (first entry)
XX
XX Listeria monocytogenes protein #199.
XX
XX Antibiacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.
XX
XX Listeria monocytogenes.
XX
XX WO200177335-A2.
XX
XX 18-OCT-2001.
XX
XX 11-APR-2001; 2001WO-FR001118.
XX
XX 11-APR-2000; 2000FR-00004629.
XX
XX (INSP) INST PASTEUR.
XX
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Faïhi H, Dehoux P,
XX Dusurget O, Chetoui F, Nedjati H, Glaeser P, Kunst F, Cossart P,
XX Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
XX Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend L,
XX Chakraborty T, Dommann E, Hain T, Berche P, Chardot A, Durant L,
XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
XX Madeno E, De Pablos B, Wehlant J, Kaerst U, Entian K, Hauf J,
XX Rose M, Voss H;
XX
XX MPI; 2002-010914/01.
XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and related
XX polypeptides.
XX
XX Claim 6; SEQ ID NO 200; 192pp; French.
XX
XX The present invention relates to the genome sequence of Listeria
XX monocytogenes EGD-e (see AB03041). The genome sequence and fragments of
XX it are useful for selecting probes and primers for detecting genes in L.
XX monocytogenes and related organisms, and for studying genetic
XX polymorphisms and other genomes. The present sequence is a protein
XX encoded by the genome sequence of the present invention. Proteins
XX expressed from the genome sequence are useful for raising specific
XX antibodies, identification of L. monocytogenes and related organisms, and
XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX B12. The genome sequence and proteins encoded by it are also useful for
XX selecting compounds that regulate gene expression and cell replication
XX and modulate L. monocytogenes-related diseases. In addition, the genome
XX sequence and proteins encoded by it are useful in pharmaceutical and
XX vaccine compositions for the treatment or prevention of infections by L.
XX monocytogenes and related organisms. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 617 AA;
XX

Query Match 28.6%; Score 957; DB 5; Length 617;
Best Local Similarity 35.1%; Pred. No. 1.9e-85;

Matches 225; Conservative 124; Mismatches 228; Indels 64; Gaps 17;
QY 1 MDHKLQRIILFDIGEDNIYAAACATRLRLVKOTDKVDROSLDDPDLKGTPTGGM 60
DB 1 MDYKIAKEILANVGSEENVRSVHCATRLRKLNVKEADKKQIESISGVSVENNAQ 60
QY 61 FQIIVG--PGVDVHV---FKEIDPATSKDIAVSTEQLDQVANNNNMWPFRAPKVLADIV 115
DB 61 LQVITIGNTVGDVYKALGSEFTKLTDGDSBIAGT---KD--SDGNFSLKADIVISGIT 114
QY 116 PLIPILVGGGGLMAINNVLAQDLFGPOSIVEMFPQISGVAMINIMASAPAPLPVLVG 175
DB 115 PILGALGSGMKGILMLITL---FG-----WLTESSGTQIILVAADSVYFLPLILA 165
QY 176 FTATKRFPGNEFLGAGIGCMAYFPTLVNGYDVAATMTAGEMPMSLFGLDVAQAGYQGV 235
DB 166 YTAARKFGANPPVAIAAAGALVYPTMILFNCGAHITFLQIP-----VILMSYSFSV 217
QY 236 LPLVAVSWLAIETEFHKRLMGTDPLITPVLTLTLGFLTFIAIGPAMRWGDLAHG 295
DB 218 IPTILAVWFSLTERPLNSKIHKAKTPLTPICMLVPLTFPLAFGLGTFISQGLASG 277
QY 296 LQGLYDFGGPVGGLFLGVYSPVITGLHQSEFPTEL-ELFNQGSFTFATASMANIAQG 354
DB 278 YTFIINLSPIVAGAFMGAFWQVLVIFGIHWGFVPIMINNLSKYGSDTMIAWGPSPNPAQA 337
QY 355 AACLAVFPLANSKELKLAGASGVSVAVGITEPAIFGVNLRMPFYIGITGAIGALT 414
DB 338 GASLGVFLKTKKPEKVALGSAALTGFEITEPSIYGTLLKXKPFVIAISGAIGAIIV 397
QY 415 ALFDIKAVLAGAGFLGVSIDAPDMV---MVLCAVVTFFVAFGAATAY-----GLYIV 466
DB 398 -----GAAGSSGAANA-IPGILTLPIFIGGFGLF--GIAVAVIISAIGTIFP 444
QY 467 RENGSIDDPATAPAPACTTYKAEAAPAFSNDSTIIQAPLTGEAIALSVSDAMFASGK 526
DB 445 ---GKDEADQI---ATTKEAKETGYE---AAVIVSPIRGINVPLNEVDENAFSAGL 494
QY 527 LGGGVAIVPTKQQLVSPVSGKIVAFPSGHAFAVTKAEDGSNVILMHIGDTYNLNGT 586
DB 495 LKGVAIVPQEGSKLISPVNGTETETAFPTGHAIGIRS----DKGVETLHVGDVTVQLNKG 550
QY 587 HFNPLKKGDEYKAGELLCEPDIDAKAGVETPTPIVSN 627
DB 551 YFKLLVAGQSDRLVQGLLEPDLAIKADGYDITPTPIVSN 591

RESULT 10
ABU32951
ID ABU32951 standard; protein; 617 AA.
XX
XX ABU32951;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by prokaryotic essential gene #18478.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Listeria monocytogenes.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX

Job time : 172 secs

PT New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.

XX Disclosure; SEQ ID NO 2958; 301bp; English.

CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPRO at
CC seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 655 AA;

Query Match 28.0%; Score 935; DB 8; Length 655;

Best Local Similarity 34.3%; Pred. No. 3.2e-83;
Matches 229; Conservative 133; Mismatches 256; Indels 50; Gaps 18;

```
QY 7 AGRILRDIGEDNIVAAACATRLRLVLEKDTKVDKQSLDDDDPLKGFETGEMFOIIVG 66
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 8 ARDLQAIGKENVATVTHCATRMRFLRDDKKANKAKESIPAVKGFITNAGQFOVILIG 67
QY 67 PGVDHVFPELDDATSKDIADVTEQKDVANNANMFSAVKVLADIFVPLIPILVGGGL 126
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 68 -NDVPIFYVDF--TAVSGIGSVSEAKTSKAKSNQNVQGVMTTLAEIFPIIPALIVGSL 125
QY 127 LMAINNVL--VAODLFGQSLVEMFPQISGVAMINLNASAPPFLPVLVGFATKRFGG 184
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 126 IIGFRNVLEGVHMSMDGKITTESQFMAGVNHFLPLPGEALFOFLPVGITWSVRKMG 185
QY 185 NEELGIGIMAMVFTLVNVDYVAATMTAGEMP--MWSLFGLDVAQAGYQTVLPVLYVS 242
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 186 SQILGIVLGLCVSPQILNAYAVASTPAADIAANWVNGYFTVNRIGYQAQVIRPALLAG 245
QY 243 WILATIEKFLHKRLMGTADELITPVLTLLLTGFLFPIAGPAMRWVDLLAHGLOGLYDF 302
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 246 LSLSYLEIFWKRKHIPVISIWFVPLSLPALIAHTVLGP---IGWTIGQLSSVLA 301
QY 303 G--GPVG---GLLFLVSPRIVYTGHSFPPIELFL--NOGGSFIFATASMANIAQGA 356
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 302 GLTGPVKWLFGLFGLVLPFVITGLHMTNALDTQLIADAGTALMPMIALSNIAQSSA 361
QY 357 CLAIFFLAKSEKLG--LAGASGVSAVIGITEPAIFGNRLRLWPFYIGIGTAIGGALIA 415
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 362 VFAYIVFMHRHDECEAQSLSLPATISAYLGTEPALFGVNVKITYIPFVAGMTGSLAGMLSV 421
QY 416 LPDIKAVALGAAGFLGVNSIDAPDVMF---LVCAVTFPVIAFGAIAIYGLYVRRNG- 470
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 422 TENVTAASIGIGLPGILSIGIQPYMLPFAGTMLVAIVVPMLLTF-----FFRKAGL 472
QY 471 --SIDDAATAAPVPAGTKAE--AEPAESNDSTIIQAPLGEATAISSVSPAMFASGK 526
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 473 FTKIBEDTNLQAEFVAQEEBAEFVSHPEVLT--SVEIISPLTGQVKELSOATDVPASGV 530
QY 527 LSGGVAIVPTKQQLVSPVSGKIIVAFPSGHAFAVTRKAEDGSNVDILMHIIGPDTVNLNGT 586
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 531 MGGGLVIEBPGGLTSPVNGTVLVLPFTKAIGI--VSDEG--VELLIHGMOTVGLDOK 586
QY 587 HENPLKKQGDVAKAGELLCEFDIDAIIKAAGYEVTTPIVVSNYK-----TGPVNTYGLGE 641
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 587 GFESLVVQGDHVIIVGOQLIRFDMDVIIKAAGLVETETVLIITNODAYTATITGYPT----T 642
QY 642 IEAGANILL 649
DB |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
 643 IQAGASLIM 650
```

Search completed: March 7, 2005, 22:11:02


```

Db      282 ATLSIPFGLGFLIGGHLIVSGVHHI FNLEVOQLADHANPENAIITAMTAQGA 341
Qy      357 CLAVFLAASEKTKLAGASGVSAVLGTEPAIFGVNLRLRMPFYIGITAAIGALLAL 416
Db      342 TVAVGVKTNPXKTLAPALASAFIGITEPAIFGVNLFRKPFSLIAGIGGLASI 401
Qy      417 FDKAVNLAGAGLVGVSIDAPDMV-----MFLVCVVTPVIAFGAALVGLVLR 468
Db      402 -----LGLAGTNGCITIIIPGTMLYVNGQLPQYLLMVAVSFALGALTTFGy----- 449
Qy      469 NSGIDPDATAPVPAGTTKAEAE-APAEFSNDSTIIQAPLTGEALISVSDAMPASGKL 527
Db      450 --EDEVDTAAAKQAEVAEKEEVAPALQNETLV--TPVGDVVALADVNDPVPSSGAM 505
Qy      528 GSGVAIVPTKQOLVSPVSGKIIVAPPSGHAFAVTRKAEADGSDVDIMHIGPTVNLNGTH 587
Db      506 GGGIAVKPSQGVVYAPADAESIAFPTGAFGLKTR---NGAEVLIVHIGDITVSMNGDG 561
Qy      588 FNPCLKQGDDEVYAGELCEFPDIDAIKAGYEVTTPVIVSN---YKKTGPVNTYGLGEIEA 644
Db      562 FEAKVAQGNKVKAGDVLGTFDSNKIIAAGLDITTMVITNTADVASVAPVAT---GSYAK 618
Qy      645 GANLINV 651
Db      619 GDVATIEV 625

RESULT 2
US-09-107-433-3250
; Sequence 3250, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Maltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinfiello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3250:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae

```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...643
; SEQUENCE DESCRIPTION: SEQ ID NO: 3250:
US-09-107-433-3250
Query Match      28.9%; Score 965; DB 4; Length 643;
Best Local Similarity 33.4%; Pred. No. 3e-90; Indels 58; Gaps 14;
Matches 223; Conservative 141; Mismatches 245;
1 MDHKDLAORIRLDIGEDNIVAAAHACATRLRLVLTQTDVDVROSIDDDPDKGPFETGGM 60
17 MNNGIARAKVIALDAGRGRENVSVAHCATRLRVVMVDEKIKVEIENLEKVGAFNNGQ 76
Qy      61 FOIIVPGEDVDHFKFE---LDDATSKDIADVSTEQLDVYANNANNMSPSAVKVLADIYPL 117
Db      77 YOIIFGTGTVMKMYDEVVVLGLPFS---SKDDMKAEVAAQGMFQRAITPFGVFPVI 131
Qy      118 IPIIVGGGLMAINNVLVAODLFPQSLVEMFPQISGVAEEMINMASAPFAFLPVVGFT 177
Db      132 IPIVATGLFNGVRGLFNALTEM-----PLPGDFATYTOITDTAFTIILPGLVWS 181
Qy      178 ATKRFGENEFLGAGIGMAVEPTLVNGYDVATATAGEMPMSLFGLDVAQAGYGTLP 237
Db      182 TFRVFGGNPAVGIVLGMVLVSGSLPNMAVA---QGGVLTAMNFFGF-IPVVGIGSVLP 237
Qy      238 VLVWSMIATTEKFLHKLMTGADFLITPVLTLLITGFLTIAIGPAMRWGDLHGLQ 297
Db      238 AFIIIGVAGKEKAVRKVPDVIDLVTVPFVTLVMSITGLFVIGPVHVENVILATK 297
Qy      298 GLYDFGVGGLFELGLVYSPVITGTHOSFPFIELEFNQGSFIF-ATASMANIAQGA 356
Db      298 AIIIGVAGKEKAVRKVPDVIDLVTVPFVTLVMSITGLFVIGPVHVENVILATK 297
Qy      357 CLAVFLAASEKTKLAGASGVSAVLGTEPAIFGVNLRLRMPFYIGITAAIGALLAL 416
Db      358 TVAVGVKTNPXKTLAPALASAFIGITEPAIFGVNLFRKPFSLIAGIGGLASI 417
Qy      417 FDKAVNLAGAGLVGVSIDAPDMV-----MFLVCVVTPVIAFGAALVGLVLR 468
Db      418 -----LGLAGTNGCITIIIPGTMLYVNGQLPQYLLMVAVSFALGALTTFGy----- 465
Qy      469 NSGIDPDATAPVPAGTTKAEAE-APAEFSNDSTIIQAPLTGEALISVSDAMPASGKL 527
Db      466 --EDEVDTAAAKQAEVAEKEEVAPALQNETLV--TPVGDVVALADVNDPVPSSGAM 521
Qy      528 GSGVAIVPTKQOLVSPVSGKIIVAPPSGHAFAVTRKAEADGSDVDIMHIGPTVNLNGTH 587
Db      522 GGGIAVKPSQGVVYAPADAESIAFPTGAFGLKTR---NGAEVLIVHIGDITVSMNGDG 577
Qy      588 FNPCLKQGDDEVYAGELCEFPDIDAIKAGYEVTTPVIVSN---YKKTGPVNTYGLGEIEA 644
Db      578 FEAKVAQGNKVKAGDVLGTFDSNKIIAAGLDITTMVITNTADVASVAPVAT---GSYAK 634
Qy      645 GANLINV 651
Db      635 GDVATIEV 641

RESULT 3
US-09-134-000C-5822
; Sequence 5822, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812

```

```

Best Local Similarity 34.3%, Pred. No. 3,96-87;
Matches 229; Conservative 133; Mismatches 226; Indels 50; Gaps 18;

QY 7 AORLRDIGEDNIVAAAHCAIRLRILVKTQDVDRSLDDDPDKGTETGSMFOIIVG 66
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 8 AKDLLOAIGGKENTYLAVTHCATRMRFYLRDCKKANVAIESIIPAVKGFYTAQGFQVILG 67
QY 67 PGVDVHFKEKLDATSDIDIAVSTQGLKDVVANNANMFSRAVKLADI FVPLPIVLNGGL 126
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 68 -NDPIEFEND- TVSGIEGVSKATKSAKSNQNVQGVTTTAAIEFTPIIPILVIGGL 125
QY 127 LMAINNVL--VAODLFGPOSIVEMFPOISGVAEIMINIMASAPFAFLPVLTGFTATKREGG 184
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 126 ILGRNVLEGHWSMLDQKITTESSQWAGVNHMLPGEAIPFQFLVGVIGTMSVRKMG 185
QY 185 NEFLAGIGNAWMEPTLVINGDYVAATMTAGEMP--MMSLFGLDVAQAQGYQTVLPVLVS 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 186 SQLIGVILGICLVSPOLMNAVAVSTPADIAANWVNNFGVFTYNRIGYQAQVILPALLAG 245
QY 243 WILATIEFLHKLMTADFLITVLTULLTGFLFTALIGPANKRWVGDLLAHGQGLYDF 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 246 LSLVYELIFMRKHIPEVISMIFVPEFLIPALIIAHTVLGP---IGMTIGQGLSSVLA 301
QY 303 G--GPGV---GLFLGLVYSPVITGTLHQSFPPIELELF--NQGSGFEFATSMANIIAOGAA 356
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 302 GLTPGVKMLRGALFALYAPFVITGLHMTNADITOLLADGGTALMPMLTASIIAOGSA 361
QY 357 CLAVEFLAKSEKLG--LAGASGVSAVIGITEPAIFGVNLRMRPFIYIGITAAIGALIA 415
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 362 VFAYFYMRHDECAQYSLPATISAYIGTEBPALFGVNVKTIYFVAGMGSLALAGMLSV 421
QY 416 LFDIKAVALGAAGFLGVSIDAPDMWF---LYCAVVTFVIAFGAALVAGLYVRRNG- 470
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 422 TFNTYTAISIGIGGLPILISIOPOYMLFPAGTMLVAIVPMLTF-----FFRKAGL 472
QY 471 --SIDPRATAVPVAGTTAAE--AEAPAEFSNSTIIQAPLTGAILMSVSDAMPKSGK 526
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 473 FTKEEGTNLQAEFVAGOEAEFVSHSEVELT--SVELIISPLTGQVKELSOATDVFASGV 530
QY 527 LGSVAIVPTKQGLVSPVSGKIIVAFPSGHAFAVRTAABDGSNDVIMHIGFDVNNGT 586
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 531 MGGGLVIEPBGSGELTSPVNGTVLTFPTKAIGI--VSDGG--VELLIHIGMDTVGLDGK 586
QY 587 HFNFLEKQGBEVKAGELLCEFDIDAIRKAGYEVTTPIVASNKK-----TGVPVYGLGE 641
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 587 GFESLVVGGHVHVGQDLIRFDMVICAAGLVETPVIITNQDAVTATITGYPT-----T 642
QY 642 IEAGANLL 649
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 643 IQAGASIM 650

RESULT 5
DS-09-107-532A-5288
Sequence 5288, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5288:
SEQUENCE CHARACTERISTICS:
LENGTH: 628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...628
SEQUENCE DESCRIPTION: SEQ ID NO: 5288:
US-09-107-532A-5288

Query Match
Best Local Similarity 32.6%; Score 923; DB 4; Length 628;
Matches 215; Conservative 134; Mismatches 253; Indels 64; Gaps 12;

QY 1 MDHKLACRIILDIGEDNIVAAACATRLRLVKDTKDVRSQSLDDPDLKGTFTGGM 60
DB 8 MGNQAVGRVWEAVAGQKNVKSIVLCATRLRFLKDESLADTQKLEPDDVIGVVGSGQ 67
QY 61 FOITVPGDVHVFEELDDATSKDIASVTEQLKDVANNANMFSRAVVKLADIPVLPFI 120
DB 68 YGVVIG-SNVADVOSIVD--EEGLAVS---EDGKESKNTINRIIDISSIFTFPLGA 120
QY 121 LVGGGLMANNVNVQDLFGPSLVEMFPQISGVLEMNLMAASAPFAFLPVLVGFTATK 180
DB 121 MAAAGVLKGFSL-----ATVAGMTLADSGTYGIIIPAAADGVFTFLPMLAFATK 171
QY 181 REGNEFLGAGIGMANVFPFLVNGYDVAATMTAGEMPWMSLFLDV--AQAQYGVTLPV 238
DB 172 KFKANQFLVALAVALMLVLP-----ATTAVAGAGEAISFGIIPVILSPSGYTSVIP 223
QY 239 LVSWILATIEKFLHKLMTADFLITPVLLTLGLFTFLAIGPAMRWGDLAHGLOG 298
DB 224 ILAVWQSKLEPFVKVIVPQFLQMLIVPLVAVVWVPLTFELAGPITVAGNALGLEN 283
QY 299 LVDFGPGVGLFGLVYSPVITGLHQSFPPIE-LEFNQGSFIFATSMANINQAAC 357
DB 284 IYGFSEYVAGLTMGSIQVFMFGHMGWVPIFMFLNIEQYGFVLMPLPILILAGGAA 343
QY 358 LAVFLAKSEKLKGLAGAGVSAVLGITEPAIFGVNLRMPFYIGITGAALGALIALF 417
DB 344 LAVALRTKTKLRALGISSTVSLRFGITEPTVYGVTLPLKPFIAACISGGIGGAIIG 403
QY 418 DIKAVA-----LGAAGFLGVNSIDAPDMVNLVCAVVTFTVAFGAALVGLVVRNSI 471
DB 404 GYKAFSSSLVSLITPTFTINTVDGVSNTVAIVATIGFVLAFTLLILGDEQODNQ 463
QY 472 IDPDATTAAPVPAAGTTAAEAPAEFENDSTIIQAPLTGAIALSSVSDMFAFGKLSGV 531
DB 464 LE-----NKHANAGBPTTSAHRTLSPLTGVLPLSEVPDGVFSSGVWGKGI 510
QY 533 AIVPTGQLVSPVSGKIVVAFPSGAFAVRTKAEDGSNDIIMHIGFDTVNLNGTHFNPL 591
DB 511 AIDPEVGEIIVAPADSGITITLPTGHAVGITT--TDGA--EILIHIGMDVVELNGNGFEIL 566

```

```

QY 592 KKQGEVAKGELLCEFDIDAIIKAAGYEVTPIVSNYKKGTPVNTVYGLGEIEAGANILNV 651
DB 567 VKQGDVYKAGDILLRFDIEAIRAAGYSVITVLTN-----TDAPRDIETL 612
QY 652 AKKEAV 657
DB 613 DQKEII 618

RESULT 6
US-09-489-039A-8212
Sequence 8212, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8212
LENGTH: 656
TYPE: prt
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8212

Query Match
Best Local Similarity 32.7%; Score 875.5; DB 4; Length 656;
Matches 216; Conservative 124; Mismatches 271; Indels 49; Gaps 13;

QY 1 MDHKLACRIILDIGEDNIVAAACATRLRLVKDTKDVRSQSLDDPDLKGTFTGGM 60
DB 37 MEYKALADILNRVVGKENIVSLVHCATRLRFLKDKNGKADAEGLKANPVMVESGQ 96
QY 61 FOITVPGDVHVFEELDDATSKDIASVTEQLKDVANNANMFSRAVVKLADIPVLPFI 118
DB 97 FQVVG--NHV-HDVLAARQEXGLSDSEPAEERKAAGSVLSQILDIISGLTFPFI 151
QY 119 PILVGGGLMANNVNVQDLFGPSLVEMFPQISGVLEMNLMAASAPFAFLPVLVGFTATK 178
DB 132 GVMATGELKGLALAAVTCMLTP-----QGYKIMFASDAPFFPFLGTTA 202
QY 179 TKRFGNEFLGAGIGMANVFPFLVNGYDVAATMTAGEMPWMSLFLDVAAQAQYGVTLPV 238
DB 203 GKFFGNFPIISWVIGALTLHPLMIOAFASQAPGA--VEHFLGIPTFTINYSSTVIP 259
QY 239 LVSWILATIEKFLHKLMTADFLITPVLLTLGLFTFLAIGPAMRWGDLAHGLOG 298
DB 260 ILASVWSCMLERKSALLPSSMKNEFSPALCLAVVPLTFVIGVATWLSHLANGYOF 319
QY 299 LVDFGPGVGLFGLVYSPVITGLHQSFPPIE-LEFNQGSFIFATSMANINQAAC 357
DB 320 IYAPFMTLAGAVLGMVOCYIFGLHMGVLPMINMTVVLGHDSLPIILRAVDAQVAV 379
QY 358 LAVFLAKSEKLKGLAGAGVSAVLGITEPAIFGVNLRMPFYIGITGAALGALIALF 417
DB 380 LGIFLATDARQVLAGSAFAGLITEPAIVGLTLPLRBPFIKGVAAIGAITAFS 439
QY 418 DIKAVALLAAGFLGVNSIDAP--DMVNF--LVCAVVTFTVAFGAALVGLVVRNSI 472
DB 440 NSYAFSGFLPNIFFPDAQMIPEGIDASVWGGLIGTVAFVLA CVLTFPAGL----- 490
QY 472 DPDATAAVVPAAGTTAAEAPAEFENDSTIIQAPLTGAIALSSVSDMFAFGKLSGV 532
DB 491 -PRGSAP-----GAVTVAPAS-AND--ILAPSGVILAEQVPDSTFASGLIGKVA 539
QY 533 IVPPTGQLVSPVSGKIVVAFPSGAFAVRTKAEDGSNDIIMHIGFDTVNLNGTHFNPL 592
DB 540 IIPAVGQVIAAPPBGVVALFQTKAIGL-----QDSGSEILIHVIGIDIVKLDGVFPFAHV 595

```



```

OY 175 GFPAITRFGNEBETLGAGIGMAAVPEPLVVGXVAAVTMTAGEMPMWSLFGJLDVDAQOQT 234
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 189 AYTBAKKENTDRITANVIAAALVPIYSAVSDTL-----RFLGMPVILAKYTSY 240
OY 235 VLEPLVWSWILATIEKFLHKLMTGADFLITPEVLTLLTGFLTEPIAIGPAMRWYDILAH 294
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 241 VIFAILAAVWLASYIEPKIRKSLHESIRNLTFPICIIYVWPIILLVVGPIADYASQILIA 300
OY 295 GLOGLVDFGEPVGULLFGLVSPVITYTGHSFPFIELELNQGSFIFANASANI-AQ 353
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 301 GYLAIVNFSVLISGAVIIGGFVOVLVIFGLHMGVWMTNNLSFYGRDITLGPACTTAAAO 360
OY 354 GAACLAFFFLAKSEKUKLAGASGVSAVIGITEPAIFGVNRLTWPFIYIGITAAIGAL 413
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 361 AGAVLIGVFLTKNKKYKXSLSLASFIALTGITPEAVYGTLLKYRPFIACTGALIGV 420
OY 414 IALFDIKAVALLGAAGFL-----GVSSIDAPDWMLVCNAVTFVIAFGAIAIAYGLYL 465
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 421 AGAAGACALVAARRSLISPFYIEGSEFWLVASYLEAMISSMFLFLFGY----- 470
OY 466 VRRNGSIDPATAAVAPVAGTTKAEAPAFENDESTIIIOAPLTGAILSLSSDAMFASG 525
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 471 -----KDEIF--EESKSIVLSTRAGEIIDLSEVNDPTFASG 506
OY 526 KLGSGVAIVETKQOLVSPVSGKIIVAFSPSGHFAVRTKAEQDSNVIDLMIHIGPTVNLG 565
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 507 SLGEGFAIITPTDGKIRYSSVNGEVSIVTFPKHAIIV--VSEEGA--EILIHIGIDTVNLG 562
OY 586 THRPPLKQODEYKAGELLCEPDIIDAKAAGEVUTPIYVSN 627
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 563 KYFQSAVSDGKVRKPDLMVEVDQELIEGSDPTMTIAVYN 604

```

```

RESULT 9
US-09-489-039A-13018
; Sequence 13018, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13018
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13018

```

	Query Match	24.3%	Score 812;	DB 4:	Length 483;
	Best Local Similarity	37.6%;	Pred. No. 1.1e-74;		
	Matches	175;	Conservative 88;	Mismatches 178;	Indels 24; Gaps 5
Oy	1	MDHKDLAORIIRDIGGEDNIVAAAHCAETRLVLTVDKTKVDHROSDDDEDLKGTEFEGM	60		
Dd	28	MDFEQISSTLLPILGSGENINASAACATRLVLVDALADQAIGIKIDGVGCFFRNAGQ	87		
Oy	61	FQIYGPGRVDHFVEKELDDATSKDLAVSTEQLDKVANNANPSRAVKLADI FVFLPI	120		
Dd	88	MOITGTGVNNRYVAFTIQAGISS-SKEAADLAGKLNPQRARLLSNIFVTIDA	146		
Oy	121	LVGGLGLMAINNVLVAODLEFPQSIVEMPEQISGAEMINLASAPFAFLPVLVGETAK	180		
Dd	147	IVASLGILMGILMWTKYGMWVDSNMLY-----MDMCSSAFLIPLIIIGTARR	197		
Oy	181	RFGNEFLCAGIGAMNVPTLVNGDIVA--TTMRAGEPMMSLRFLDVAQAGYGCTVP	237		
Dd	198	EFGGNPYTGATLGGILLTHPALTNMGVAAAGFHIM-----NFQIEVAMIGYGQCTVP	249		

[illegible]

```

RESULT 10
US-09-543-681A-7019
; Sequence 7019, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7019
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7019

```

Query Match	Similarity	24.0%	Score	802.5	DB.4	Length	634
Best Local	21.1	32.0%	Pred.	NO. 1.8e-73			
Match	211	Conservative	129	Mismatches	231	Indels	89
						Gaps	22
QY	1	MDHKDLAQRILRDIGGEDNIVAAAHGATRLRLVLTQTKDVRQSLDDPDGLKGTETGGM	60				
DB	6	MNYKEFAKQILAIYVGGEDNIKSLVHCSTRLEFTHNEDKIPSPKAKANSKILSDVMKGGQ	65				
QY	61	FOIIVGPDVQVHVEFLDDATSKDIAVST-----EDLKDVANNAMWFSRAVVLADI	113				
DB	66	YQLYVG-NDVQVFNEL-----TKMKITNTPHDNNKDKQKGSIVN-----KILSAI	111				
QY	114	---FVFLPIPIVGGGLMIAINNVLVAODLFGQSLVEMFPOISGVAEMINLMASAPPAFL	170				
DB	112	TGSLAPVLPPLAAGMGKVLIVL-----TALGLMLDKSDQTYUILLNFENFAFYEM	162				
QY	171	PVLVGFTATKEFGNGEFLGAGIGMAMVPEPTLVNGDVATMAGEMPWMSLGLDVAQAG	230				
DB	163	PVLTGFSAKTFNCNQYLAAVCCLMASVP-----EMTSLVAKG-PV-EFHFIYALVAK	214				
QY	231	YQGTLPVLVYVSWILATIEKFLAHRKLMGTADFLITPVITLLLTGFLTPIAGPAMRWYGD	290				
DB	215	YSSQLIPALITVMMSYTERFIVRIIVPEMWKVFMPVLVIVSTPIALIAVGPVTSWFAQ	274				
QY	291	LLAHLQGLNDPGEVGGLLPGLVYSPVITGLHOSFPIELTELPNQG--SFIFATASM	348				
DB	275	LIADSVILIQHTGTITAIPLVAIYPMVLVSGMHRALSPVSIIMLVQKQFDDIRVMALC	334				
QY	349	ANIAGAACLAVFLPAKSEKLGLAGASGVSAVL-GITTEPALFGVNLRLMWFYIGIGTA	407				
DB	335	SNMSGAAASLAVSVYTKKTKTLQGLAFSASITFAFGGITPEPAMYGVNLKLKEMTY-----AC	390				
QY	408	ALGGLALMF-----DIKAVLGAAGFLV---VSTADADMVFLVCAVATPPIAAGAIA	460				
DB	391	MGGGAIAGLFGAIVLKAIFYVYTPBLBIPMWSI-DTDNQV---UNAIITTLIASVATRI	446				

QY 144 SLVEMPQISGVAEINILMASAPFAFLPYLVGTLATKREGNEPLGAGIGMAMVPTLVN 203
 DB 65 TIVESHKFNKGNVNAFIMLPGEALFHPFVGITWISIAKKMGTTQIILGIVLITVBPOLN 124
 QY 204 GYDVATATMAGEMPMWSLFGLDVAQAQYOGTVLPVLVSMILATIEKPLHKRLMGTADEL 263
 DB 125 AYSVAST-AAADIPFDFGPAQVOMGYQAQVLPALAGFMALYEIFPRKXIPQGISMI 183
 QY 264 ITFVTLTLTGFTLFAIGP---AMRWGDLIAHLOGLYDFGGVGLGLGYSPV 319
 DB 184 FVPLFSLPVLAAHYLTSPIGWTGSMISNYNCTLTSSIXW---LFSAVGFLAPLV 240
 QY 320 ITGLHOSPPETIELEFNQGSF-IFATASMANIAOGACLAFLAK-SEKXKGLAGAG 377
 DB 241 ITGLHMANNAIDMQLIADGSGTUMPMILSNIAQSAYALITFLHKKKEEQISIPM 300
 QY 378 VSAVLGITEPAIFGVNLRMPFYIGITPAIGALIALFDIKAVLAGAAGLVGVSIDA 437
 DB 301 ISCYGVTEPAMGINKLVYFPVAMVSGLAGMVRANLIGVGLPGLIAQA 360
 QY 438 PDMWFLVCAVTVTVFVAFGAALAYGLVYVRNG---SIDP-DATAAPVPAITKAABA 492
 DB 361 ETWVPEIIMIAVILPFGILTIIF---RQGIINKIDPAVPENADVQLQTANGATAT 415
 QY 493 PAEFSNDSTIIQAPLTGEALA-----LSSVSDAMFASGLSGVAIVPTK 537
 DB 416 POSFEVSA-----TGNAVATKETLFAVAAGTIKTEVNDVFSQKMGQGVAVEPSN 469
 QY 538 GOLVSPVSGKIVAFPSGHAFAVRTKAEQSDNVDILMHIGFTVNLNTHFNPLKKQGE 597
 DB 470 GKYYAVNKGVTSEVFTKHAIGILS---NEGLEVLVHMGDLTVLKGVPFVYVGEVL 525
 QY 598 VYAGELCEPDIDALKAAGYEVTTPVIVSNYKKTGVNTYTGGEIAG 645
 DB 526 VPEETLIAEWDLPEIQAGKTDIIVALTNNKEVAGLSIDQSLVRPG 573
 RESULT 13
 US-09-107-433-5169
 / Sequence 5169, Application US/09107433
 / Patent No. 6800744
 / GENERAL INFORMATION:
 / APPLICANT: Lynn A Doucette-Stamm and David Bush
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
 / SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
 / THERAPEUTICS
 / NUMBER OF SEQUENCES: 5206
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 / STREET: 100 Beaver Street
 / CITY: Waltham
 / STATE: Massachusetts
 / COUNTRY: USA
 / ZIP: 02354
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: CD-ROM ISO9660
 / SOFTWARE: <Unknown>
 / OPERATING SYSTEM: <Unknown>
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/107,433
 / FILING DATE: 30-Jun-1998
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 60/ 085131
 / FILING DATE: May 12, 1998
 / APPLICATION NUMBER: 60/051553
 / FILING DATE: July 2, 1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Ariniello, Pamela Deneke
 / REGISTRATION NUMBER: 40,489
 / REFERENCE/DOCKET NUMBER: GTC-011
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (781) 893-5007

; TELEFAX: (781) 893-8277
 ; INFORMATION FOR SEQ ID NO: 5169:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 530 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Streptococcus pneumoniae
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...530
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5169:
 US-09-107-433-5169
 Query Match 20.3%; Score 678; DB 4; Length 530;
 Best Local Similarity 33.7%; Pred. No. 9.2e-61;
 Matches 168; Conservative 101; Mismatches 184; Indels 46; Gaps 15;
 QY 174 VGFATKFGGNEFGAGIGMAMVPTLVNGYDVAAITTAGEMP--MMSLGLDVAQAQY 231
 DB 50 ITWSVRKQGTNQLIGVIGLVSPPOLLNVAVASTEPADIAANWVNFQFTVNRIGY 109
 QY 232 OGTVLPVLVSMILATIEKPLHKRLMGTADELITPVLTLTLTGFTLFAIGPAMRWGDL 291
 DB 110 QAOVLPALLAGLSYLEIFPRKPIPEVISMIFVDFSLIPALLIAHTVLGP---IGWT 165
 QY 292 LAHQGLQGYDFG--GPVG---GLLFGVYSPVITGTHQSPPEIELEF--NOGGSFIPAT 345
 DB 166 IQGGLSVVLAGLTGPVFMFLFGAIGALYAPVITGLHMTNAIDTQLIADAGGTALWPM 225
 QY 346 ASMANIAOGACLAIFPLAKSEKLG--LAGSGVAVNGITEPAIFGVNLRMPFYIGI 404
 DB 226 IALSNIAGQSAVFAYFMRHDECAQVSLPATTISAVYGVTEPAIFGVNXXIYFVAGM 285
 QY 405 GTAAGCALIALFDIKAVLAGAAGLVGVSIDAPDMNF---LYCAVTFVIAFGAIA 460
 DB 286 TGSALAGMLSTFNTAASIGIGLPGILSTOPQYMLFPAGTMLVAYVPMILTT----- 340
 QY 461 YGLVYVRNG---SIDPDATAAPVPAITKAEE--AEAPAEFSNDSTIIQAPLTGEALA 515
 DB 341 ----PFRKAGLPETKIEGDTNLQAEFVAQBEAFVSHBEVELT--SVEIISPLTGQVKE 394
 QY 516 SVSDAMPASGKLGSGVALVPTKQGLVSPVSGKIVAFPSGHAFAVRTKAEQSDNVDILMH 575
 DB 395 QATDVPFASGVWGQGLVIEPSQGLTSPVNGTVTLFPTKHAIGI--VSDG--VELLIH 450
 QY 576 IGFDTVNINGTFFNPLKKQGEVKAGELCEPDIDALKAAGYEVTTPVIVSNYK----- 630
 DB 451 IGMDTVGLDGKGFESLVQGDHVIYGOQLIRFDMVITKAGLVETPVIITNQDAYTATI 510
 QY 631 TGPVNTYGLGEIEAGANLL 649
 DB 511 TGTPT-----TIQAGASLM 525
 RESULT 14
 US-09-543-681A-5156
 / Sequence 5156, Application US/09543681A
 / Patent No. 6605709
 / GENERAL INFORMATION:
 / APPLICANT: GARY BRETON
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
 / FILE REFERENCE: 2709.1002-001
 / CURRENT APPLICATION NUMBER: US/09/543,681A
 / PRIOR FILING DATE: 2000-04-05
 / PRIOR APPLICATION NUMBER: US 60/128,706
 / NUMBER OF SEQ ID NOS: 8344
 / SEQ ID NO 5156
 / LENGTH: 481

```

; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5156

Query Match      19.5%; Score 650.5; DB 4; Length 481;
Best Local Similarity 31.7%; Pred. No. 5.4e-58;
Matches 150; Conservative 111; Mismatches 197; Indels 15; Gaps 10;

QY      8 QRLRDIGEDNIVAAAHGATRLRLVLDKTDVDRSLDDPDLKGTFTGGMFQIIVG 67
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      19 EELITLVGSDNIVSYSHCLTRLRFLVLPNPLADIDIKIKISFVKCFNNAGQFQYIVGM 78
QY      68 GVDVHFKELDAKTSQDIATSTEQLDKDVANNANWFSRAVKYADI FVPLPILVGGGL 127
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      79 -DVIDIYKLLIDKLSK-AKSKDIDKSDAKONGSQKLIANLAEI FVPLPILVGGGL 136
QY      128 MAINVLVAODLFGPQSLVEMFPQISGVAEMINLMASAPAPLPVLVGFATRRFGNGER 187
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      137 LGLRNVIGEMPISDNKPRTFYFPMFLPIYDFLWLPCEALFHFPLVSGI CMSSVKKMGSTA 196
QY      188 LGAGIGMANVFPVLVNGYDVAATMTAGEMP-MWSLFLG-DVAQAGTGYVLPVLVSWIL 245
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      197 LGLVLDITIVSPQIMANVNGS-----QLPEVWN-FGLFTIEKVGQYQACVIPSILGLFL 250
QY      246 ATIEKFLHKLMTADFLITPVLTLLTGFLTIAIGPMRWGDLAHGLGLY--DYG 303
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      251 GMEFTKRRPIPYLKLVIPIVTLVAVFAHVILGPIGRSIGDGLAQLIKPLMGDP- 309
QY      304 GPVGLLFGLVSPYITVITGLHOSFPPIELFL-NOGGSFIFATASNANIQAACLAVPF 362
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      310 APIGSTIFAFYSPWITGHTHTLAIIDQMOSYGTPIWPIALISNIAQASAVVGIII 369
QY      363 LAKSEKLGAGAGSVAVLIGTEPAIFGVNLRNRPFIYIGTAAIGALLAIPDIKAV 422
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      370 VSSKLNQEOETITPAISAYLGVTEPPAMYGNLKYGFPMLCAMIGALLAGICLNINVLN 429
QY      423 ALGAGFLGVSTIDAPDMWFLVCAVTFPIARGAALVGLVLRNGSIDPD 475
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      430 GIGVGLPGILSIQPTYVLTLMALVAVPIGLTILVRY-KERAGTLQVD 481

RESULT 15
US-09-107-532A-6004
; Sequence 6004, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneka
; REGISTRATION NUMBER: 40,489

```

```

; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6004:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (b) LOCATION 1...590
; SEQUENCE DESCRIPTION: SEQ ID NO: 6004:
US-09-107-532A-6004

Query Match      19.3%; Score 644; DB 4; Length 590;
Best Local Similarity 26.6%; Pred. No. 3.6e-57;
Matches 164; Conservative 130; Mismatches 249; Indels 74; Gaps 11;

QY      3 HKDLAQRILRDIGEDNIVAAAHGATRLRLVLDKTDVDRSLDDPDLKGTFTGGMFQ 62
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      24 YELARKIVENGVKENINGLTHCTTRLRFLKNEBKANDELKMDGIVTMRAGGYQ 83
QY      63 IIVGPDVDHFKELDDA--TSKDIASVSTEQLDKDVANNANWFSRAVKYADI FVPLPIL 120
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      84 VVIG-NHVVVEEYIKAGNLTFDEAVSTKKRP-----FDMLIDISCFPLAI 134
QY      121 LVGGGLMANVVAODLFGPQSLVEMFPQISGVAEMINLMASAPAPLPVLVGFATRK 180
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      135 LAAGMIRGLTFLV-----FLGAPDRSGTFVFDNIGDSVFQFMPVIGITAA 185
QY      181 RFGNEFLGAGGMVFPVLVNGYDVAATMTAGEMPMSLFGLDVAQ----- 229
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      186 KKVNEFVGMILGAALMNSL-----SLELSGAAEAPLTITISGTFEAPITYQVPIGW 241
QY      230 ---GYQTVLPVLVSWILATIEKFLHKLMTADFLITPVLTLLTGFLTIAIGPMRW 286
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      242 IARNVASSVPIIFIVLASQIQKPIKLVPEMINAFVFPFTVLITWPLGLVGFVPT 301
QY      287 WTGDLIAHGLQGYDPCGPGVGLLFGLVSPYITVITGLHOSFPPIELFLN-OGGSFTAT 345
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      302 FATDILMAGFETILASPVIYGAIVGFQIIVMFGHLHVAIYPMGIMOPSVNGMONIMTP 361
QY      346 ASMANIQAACLAFLAKSEKLGAGAGSVAVLIGTEPAIFGVNLRNRPFIYIGT 405
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      362 VAVVSFGQALPALTALYFKLRNPKDAIAPAVSGIVGITEPAIYGTLPKRIPIFTCV 421
QY      406 TAAIGGALLALPDIKAVLAGAAGFLGVSI-----DAPDMWFLVCAVTFPIARGAALA 460
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      422 GGALGAGVSGLMNLSMNGGGLGIFTIPYIRPDGDLTDVINVLGIALAMVSVFLTFP 481
QY      461 YGLVYVRNGSIDPDATAPVPAGTTKAEAEAPAEFSNDS--TIQAPLTGEALISVS 518
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      482 F-----WKDEAGETDIOKKSKEIYKPTIOGIAPLNAAK 517
QY      519 DAMPASGKLGSAVAVPTPKGOLVSPVSGKIIVAFPSGNAFAVTRTADOSANDVIMHIF 578
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      518 DAFAQGTIGRGILLYPEKGEVRAFPDGTWTLFTKQALGVSR---TGLELIHVL 573
QY      579 DTVINNGTHFNPLKQG 595
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      574 DTVOLEGRKFESLVQGG 590

Search completed: March 7, 2005, 21:57:09
Job time : 46 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:54:07 ; Search time 136 Seconds
(without alignments)
1599.255 Million cell updates/sec

Title: US-10-019-284B-2

Perfect score: 3342

Sequence: 1 MDHDLARIRLDIGEDNI.....IEAGNLNVAKKEAVPATP 661

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3310	99.0	661	US-09-738-626-6404	Sequence 6404, App
2	2332	69.8	468	US-10-450-055-10	Sequence 10, App1
3	957	28.6	617	US-10-282-122A-60875	Sequence 60875, A
4	956	28.6	627	US-09-815-242-13467	Sequence 13467, A
5	949	28.4	655	US-10-474-776-375	Sequence 375, App
6	949	28.4	655	US-10-474-928-3912	Sequence 3912, App
7	801.5	24.0	612	US-10-282-122A-73843	Sequence 73843, A
8	799.5	23.9	612	US-10-474-776-253	Sequence 253, App
9	799.5	23.9	612	US-10-472-928-1038	Sequence 1038, A
10	785.5	23.5	620	US-10-282-122A-74439	Sequence 74439, A
11	738	22.1	379	US-10-282-122A-70209	Sequence 70209, A
12	725.5	21.7	683	US-09-738-626-6961	Sequence 6961, App
13	725.5	21.7	683	US-10-450-055-26	Sequence 26, App1

14	627	18.8	451	US-10-282-122A-57628	Sequence 57628, A
15	584	17.5	381	US-10-474-776-367	Sequence 367, App
16	554	16.6	453	US-10-282-122A-75796	Sequence 75796, A
17	550	16.5	484	US-09-815-242-10809	Sequence 10809, A
18	550	16.5	484	US-10-282-122A-56917	Sequence 56917, A
19	548.5	16.4	429	US-10-282-122A-73368	Sequence 73368, A
20	538	16.1	455	US-09-815-242-13794	Sequence 13794, A
21	531	15.9	453	US-10-282-122A-55853	Sequence 55853, A
22	529	15.8	334	US-10-472-928-3552	Sequence 3552, App
23	522	15.6	484	US-09-815-242-12272	Sequence 12272, A
24	517	15.5	484	US-10-282-122A-44138	Sequence 44138, A
25	517	15.5	484	US-10-470-0488-229	Sequence 229, App
26	511	15.3	474	US-10-282-122A-70588	Sequence 70588, A
27	510	15.3	454	US-10-282-122A-45789	Sequence 45789, A
28	510	15.3	454	US-10-282-122A-68601	Sequence 68601, A
29	507	15.2	466	US-10-282-122A-53222	Sequence 53222, A
30	496	14.8	526	US-10-282-122A-57960	Sequence 57960, A
31	495.5	14.8	450	US-10-282-122A-71596	Sequence 71596, A
32	490.5	14.7	439	US-09-815-242-5660	Sequence 5660, App
33	481	14.4	403	US-10-282-122A-52840	Sequence 52840, A
34	472	14.1	545	US-10-156-761-11459	Sequence 11459, A
35	464	13.9	454	US-10-282-122A-45636	Sequence 45636, A
36	462.5	13.8	480	US-10-282-122A-60159	Sequence 60159, A
37	450.5	13.5	476	US-10-282-122A-67341	Sequence 67341, A
38	438	13.1	474	US-09-815-242-10243	Sequence 10243, A
39	438	13.1	474	US-10-282-122A-56632	Sequence 56632, A
40	437	13.1	482	US-10-282-122A-76283	Sequence 76283, A
41	436.5	13.1	485	US-10-282-122A-43101	Sequence 43101, A
42	434	13.0	483	US-10-282-122A-56083	Sequence 56083, A
43	418.5	12.5	477	US-10-282-122A-60299	Sequence 60299, A
44	374	11.2	249	US-10-474-776-368	Sequence 368, App
45	374	11.2	249	US-10-472-928-3554	Sequence 3554, App

ALIGNMENTS

RESULT 1
US-09-738-626-6404
Sequence 6404, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6404
LENGTH: 661
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6404
Query Match 99.0%; Score 3310; DB 9; Length 661;
Best Local Similarity 98.9%; Pred. No. 2.1e-275;
Matches 654; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

QY      1 MDHKLQALRILFDIGEDNITVAACHATRLRLVLTQKTVDFQSJDDDDPKGTETGGM 60
Db      1 MDHKLQALRILFDIGEDNITVAACHATRLRLVLTQKTVDFQSJDDDDPKGTETGGM 60
QY      61 FOIIVGPDVHVFKELDATSKDIASVTEOLKDVANNANMFSAVAVLADI FVPLLP I 120
Db      61 FOIIVGPDVHVFKELDATSKDIASVTEOLKDVANNANMFSAVAVLADI FVPLLP I 120
QY      121 LVGGGLLMAINNVLVAODLFGPQSLVEMFPQISGVAEMINIMASAPFAFLPVLTGFTATK 180
Db      121 LVGGGLLMAINNVLVAODLFGPQSLVEMFPQISGVAEMINIMASAPFAFLPVLTGFTATK 180
QY      121 LVGGGLLMAINNVLVAODLFGPQSLVEMFPQISGVAEMINIMASAPFAFLPVLTGFTATK 180
Db      121 LVGGGLLMAINNVLVAODLFGPQSLVEMFPQISGVAEMINIMASAPFAFLPVLTGFTATK 180
QY      181 RFGNEPFGAGIGMAMVPPSLVNGYDVAATMAAGMPMWSLFGLDVAAGYGTLPVLV 240
Db      181 RFGNEPFGAGIGMAMVPPSLVNGYDVAATMAAGMPMWSLFGLDVAAGYGTLPVLV 240
QY      241 VSMILATIEKFLHKLMTADFLINPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLY 300
Db      241 VSMILATIEKFLHKLMTADFLINPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLY 300
QY      301 DFGGVGGLLFGLVSPVITGLHQSPPPIELELFNQGSSFIPTASMANIAQGAACLA 360
Db      301 DFGGVGGLLFGLVSPVITGLHQSPPPIELELFNQGSSFIPTASMANIAQGAACLA 360
QY      361 FFLASEKLGKLAGASGVSAVLGTEPAIFGVNLELRMPFYIGITAAIGALLLPDIX 420
Db      361 FFLASEKLGKLAGASGVSAVLGTEPAIFGVNLELRMPFYIGITAAIGALLLPDIX 420
QY      421 AVALGAAGFLGVASIDAPDMWFLVCAVTEFVIAFGAIAAGLYLVRRNGSIDPATAP 480
Db      421 AVALGAAGFLGVASIDAPDMWFLVCAVTEFVIAFGAIAAGLYLVRRNGSIDPATAP 480
QY      481 VPAGTTKAEAEAPAEFSNDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQL 540
Db      481 VPAGTTKAEAEAPAEFSNDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQL 540
QY      541 VSPVSGKTVVAPSPSHAPVTRKADGSVDLIMHIGFDVNLNTHNPPLKKQDEYKA 600
Db      541 VSPVSGKTVVAPSPSHAPVTRKADGSVDLIMHIGFDVNLNTHNPPLKKQDEYKA 600
QY      601 GELLCEFDIDAIIKAAGEVETTPVIVSNYKKTGPVNTYGLGEIEAGANLINVAKKEAVPAT 660
Db      601 GELLCEFDIDAIIKAAGEVETTPVIVSNYKKTGPVNTYGLGEIEAGANLINVAKKEAVPAT 660
QY      661 P 661
Db      661 P 661

RESULT 2
US-10-450-055-10
; Sequence 10, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft
; TITLE OF INVENTION: No. US20040043953A1el genes of Corynebacterium
; FILE REFERENCE: 936 2000
; CURRENT APPLICATION NUMBER: US/10/450,055
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 10
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-450-055-10

Query Match      69.8%; Score 2332; DB 15; Length 468;
Best Local Similarity 98.5%; Pred. No. 36-192;
Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY      194 MAMVPTLVNGYDVAATWTAGEMPWMSLFGLDVAAGYGTLPVLVSMILATIEKFLH 253

```

```

Db      1 MAMVPTLVNGYDVAATWTAGEMPWMSLFGLDVAAGYGTLPVLVSMILATIEKFLH 60
QY      254 KRLMGTADFLITTPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLYDFGSPVGGILFGL 313
Db      61 KRLMGTADFLITTPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLYDFGSPVGGILFGL 120
QY      314 VYSPVITGLHQSPPPIELEFNQGSFIPTASMANIAQGAACLA VFLAKSEKLGIA 373
Db      121 VYSPVITGLHQSPPPIELEFNQGSFIPTASMANIAQGAACLA VFLAKSEKLGIA 180
QY      374 GASGVSAVLGITEPAIFGVNLELRMPFYIGITAAIGALLLPDIXKALVGAAGFLGV 433
Db      181 GASGVSAVLGITEPAIFGVNLELRMPFYIGITAAIGALLLPDIXKALVGAAGFLGV 240
QY      434 SIDAPDMWFLVCAVTEFVIAFGAIAAGLYLVRRNGSIDPATAPVPAGTTKAEAP 493
Db      241 SIDAPDMWFLVCAVTEFVIAFGAIAAGLYLVRRNGSIDPATAPVPAGTTKAEAP 300
QY      494 AEFNSDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQLVSPVSGKTVVAP 553
Db      301 AEFNSDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQLVSPVSGKTVVAP 360
QY      554 SGHAPVTRKADGSNDLIMHIGFDVNLNTHNPPLKKQDEYKAGELICEFDIDAIIK 613
Db      361 SGHAPVTRKADGSNDLIMHIGFDVNLNTHNPPLKKQDEYKAGELICEFDIDAIIK 420
QY      614 AAGEVETTPVIVSNYKKTGPVNTYGLGEIEAGANLINVAKKEAVPATP 661
Db      421 AAGEVETTPVIVSNYKKTGPVNTYGLGEIEAGANLINVAKKEAVPATP 468

RESULT 3
US-10-282-122A-60875
; Sequence 60875, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykund, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

```



```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 60875
; LENGTH: 617
; TYPE: PR
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60875

Query Match      28.6%; Score 957; DB 15; Length 617;
Best Local Similarity 35.1%; Pred. No. 1.8e-73;
Matches 225; Conservative 124; Mismatches 228; Indels 64; Gaps 17;

QY 1 MDHKLDAORILRDIGEDENIVAAAHCAATRLRLVLTQTKVDQSDDDPDLKGTFTGGM 60
DB 1 MDYQIAKSEILANVGSEENVRSVHCATRLRFLVKKERADKQISIGSVENNAQ 60
QY 61 FOIIIVG--PGVDVHV---FKEIDDATSKDIAVSTEQDKDVANNANWFSRAVKVLADIFV 115
DB 61 LQVITIGNTVGDVYKALGFTKLTDGDSIAKGT---KD---SDGNFLSKAIDVISGIFT 114
QY 116 PLIPILVGGGLMANNVLAODLFGPQSLVEMFPOISGVAEMINIMASAPAFPLVAVG 175
DB 115 PILGALAGGGMKGLMLITLTT---FG-----WLTSSTGYOILYAADSVFVFLPLILA 165
QY 176 FFAATRFQGNNEFLGAGIGMAVFPPLVNGVDVAATMTAGEMPMWSLFGLDVAQAGYQTV 235
DB 166 YPAARFEGNPPVPAALAAAGALVPTMINLFNENCAHTFLQIP-----VLMSTYSFSV 217
QY 236 LPLVNSMILATIEKFLHKLMTADPLITPVLTLLTGFLFPIAIGPAMRWGDLAAG 295
DB 218 IITIIIAVWFLSLIERLNKIHHAATPLTPMCLMLVPLTFLAAGPLGTPISGLAAG 277
QY 296 LQGLYDFGPGVGLLFGLVYSPVITGLHQSPPPIEL-ELFNQGSFIFATASMANIAG 354
DB 278 YTFIVNLSPVAVAGAFGAWQVLVIFGIHWGFVPIMINLSRGSDTMIAMVGSNFAQA 337
QY 355 AACLAVFLAKSEKLGKLAGASGVSALGITTEPAIGVNLRLRMPFYITIGTAAIGALI 414
DB 338 GMSLGFELTKKPEVVALIAGSALATGFTGITEPSIYGLTKKPPVSIAGATGALIV 397
QY 415 ALFDIKAVALGAAGFLGVNSIDAPDWV---MFLVCAVVTFFVIAFGAALAY-----GLYLV 466
DB 398 -----GAAGSSGANA-IPGLITLPIFGKPFVGL--GIAVAYIIISALGTTF 444
QY 467 RRGNSIDPDATAVPVAGTTTAAEAAPAFSNDSTIIQAPLTGEAIALSSVSDAMFASG 526
DB 445 ---GYKDEWADGI---APTTEAKETGVE---AEVIVSPIRGNIVPLNEVDEAFSAGL 494
QY 527 LGSVAIVPTKQOLVSPVSGKIVVAFPSGHAFAVRTKADGSDNDILMHIGPDTVNLGT 586
DB 495 LKQVAIVVQEQEKLISPVNGITETAPFTGHAIGIRS---DKGVELLHVGPDTVOLNKG 550
QY 587 HFNPLKQGDDEVKAGELCEFDIDAIAKAGYEVTTPVIVSN 627
DB 551 YFKLVAGQDRVLVGOALLFEDLEAIKAGDITTPVIVTN 591

RESULT 4
US-09-815-242-13467
; Sequence 13467, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cair, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes
```

```

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13467
; LENGTH: 627
; TYPE: PR
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13467

Query Match      28.6%; Score 956; DB 9; Length 627;
Best Local Similarity 33.3%; Pred. No. 2.3e-73;
Matches 222; Conservative 141; Mismatches 246; Indels 58; Gaps 14;

QY 1 MDHKLDAORILRDIGEDENIVAAAHCAATRLRLVLTQTKVDQSDDDPDLKGTFTGGM 60
DB 1 MNNQIAKVIDALDRENVASVHCAITRLRMVDEEKINKEVENLEKVGAFNSQ 60
QY 61 FOIIIVGPGVDVHVEK---LDDATSKDIAVSTEQDKDVANNANWFSRAVKVLADI FVPL 117
DB 61 YQIIFGTGVNKMVDEVVVLGLPTS-----SKDDKAEVAKQGNFQRAIRTFGDVFPBI 115
QY 118 IPIIVGGGLMANNVLAODLFGPQSLVEMFPOISGVAEMINIMASAPAFPLVAVG 177
DB 116 IPVIAVATGFMGVRSLFNALFV-----PLRDFEYTYQILTDPAFIILPGLVWS 165
QY 178 ATKRFQGNNEFLGAGIGMAVFPPLVNGVDVAATMTAGEMPMWSLFGLDVAQAGYQTV 237
DB 166 TFRVFGNPAVGIYGMMLVSGSLPNANAVA---OGGEVTAMNFGF-IPVGLQGSVLP 221
QY 238 VLVNSMILATIEKFLHKLMTADPLITPVLTLLTGFLFPIAIGPAMRWGDLAAGIQ 297
DB 222 AVIIGVAKPEKAVRKVVPDVITDLVTPVTLVMSITLGLFVIGPVFHVENVILIAVK 281
QY 296 GLYDFGPGVGLLFGLVYSPVITGLHQSPPPIEL-ELFNQGSFIF-ATASMANIAGQA 356
DB 282 AILSNPFGIGFLIGVQLIVSGVHHI FNLLEQQLAADANFPNIIITAAATACGA 341
QY 357 CLAIVFLAKSEKLGKLAGASGVSALGITTEPAIFGVNRLRMPFYITIGTAAIGALI 416
DB 342 TVAVGVTKNPKLTALPAALSAFLGITTEPAIFGVNRFKRPFLSLIAGIAGGLASI 401
QY 417 FDIKAVALGAAGFLGVNSIDAPDWV-----MFLVCAVVTFFVIAFGAALAYLYVR 468
DB 402 -----LGLATGNGITIIIPGTMVYNGQLPQYLLWVAASFALGFALTVMFGV----- 449
QY 469 NSCIDPDATAVPVAGTTTAAEA-APAFSNDSTIIQAPLTGEAIALSSVSDAMFASG 527
DB 450 --EDVEDATAAKQAEVABEKEEVAPALQNETLV--TPVIGDVVALLADVNDPVFSSGM 505
QY 528 GSGVAIVPTKQOLVSPVSGKIVVAFPSGHAFAVRTKADGSDNDILMHIGPDTVNLGNH 587
DB 506 GGGIAVKSQGVVYALADAVSIAPFTGHAIGLTKR---NGAEVLHVGIDTVSMNBDG 561
QY 588 FNPCLKQGDDEVKAGELCEFDIDAIAKAGYEVTTPVIVSN---YKKTGPVNTYGGIEA 644
DB 569 FEAQVAGQNKVYAGGVLTGTPSNKIAAAGGLDOTTNVIYTNADVASVAPVAT---GSYSK 618
```

QY 645 GANLINV 651
Db 619 GDVAVIEV 625

RESULT 5

US-10-474-776-375
; Sequence 375, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYX
; FILE REFERENCE: AM100649-PCT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 375
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-375

Query Match 28.4%; Score 949; DB 16; Length 655;
Best Local Similarity 34.9%; Pred. No. 9.8e-73;
Matches 234; Conservative 128; Mismatches 253; Indels 56; Gaps 19;

QY 7 AQRILRDIGEDNIVAAAHCAITRLRLVLTCDTVDKROSLDDPDLKGFTEGMPQIIIVG 66
Db 8 AKDLLOAIGKENYAVVTHCATRMRFVLGDDKKANVKAIESIPAVKGFITNAGQFOVILIG 67
QY 67 PGDVHVFKELDATSKDIIVSTEQIKDVANNANWFSRAKVYIADIFVPLIPIIVGGGL 126
Db 68 -NDVPIFYNDP-TAVSGIEGVSKKAASAKSNQNVGVGWTTLTAIEFTPIIPALIVGL 125
QY 127 LMAINNVL--VAODLFGPQSLVEMFPQISGVAMINIMASAPFAFLPVLVGTATKRRFG 184
Db 126 ILGFNVLEGVWMSLDKTTESQFVAGVNHFLMLPGEALFQFLPVGITMSVSRKMG 185
QY 185 NEFLGAGIGMANVPFTLVNGVDVAATMTAGEMP--MMSLFGDLVDAQYQGTLPVLVVS 242
Db 186 SQTIGIVIGICLVSPQLNAAVASTPADIAANWVNGFYTVNRIGYQAOVIRPALLAG 245
QY 243 WILATIEKRLHRLMGTAFLITPVLTLLTGFLTFIAGPAMRWGDLAHGLOGLYDF 302
Db 246 LSLSTLEIFMWHKRIEIVISMIFVPLSLIPALILAHVTLGP---IGWTIGQGLSSVLA 301
QY 303 G--GPVG---GLFGLVSPVITVIGLHOSFPPIELELF--NOGGSFIPTASMANIAGAA 356
Db 302 GLTGPVKMLFGAIFGALVAPVITIGLHMTNADPOLADAGGTALMPMIALSNIAQSA 361
QY 357 CLAVFELAK--SEKGLAGAGSVAVLGTTEPAIFGVNLRRLRPYIGIGTAAGALIA 415
Db 362 VFAVYFMHDEREQAQVSLPATISAVLGTTEPALFGVNVKTYIPFVAGMTGALAGMLSV 421
QY 416 LEDIKAVNALGAAGFLGVNSIDAPDMWF---LYCAVVTFFVIAFGAALVGLYVRRNGS 471
Db 422 TTNVTAASIGIGLPGILISIQOYMLPFAGTMLVAIVPMLTLF-----PFRKAGL 472
QY 472 IDPDATAVPVAGTTKAAE---APAEFSNDSTI-----IOAPLTGEAIALSSVSDAM 521
Db 473 FTK-----TEGDTNLQAEFVAQBEAFVNHEPVELTSVEIISPLTQVKELSQATDPI 525
QY 522 PASGKLGSVAIVPTKQGLVSPVSGKIVVAPPSGHAFAVTRKAEQDSNVDILMHIGPDTV 581
Db 526 PASGVMGGGLVIEPSQGLTSPVNGTVVLPFTKHAIGI--VSDRG--VELLIHIGMDIV 581
QY 582 NNGTHFNLKXKQSGEVKAGELCEPDIDAIKAAGEVTPPIVVS---YKKTGPVNTYG 638
Db 582 GLDGGKFBESLVVQGDHVTVGQOLIRFDMVIRKAGLVETTPVITINQADAYTATIP-GTYP 640
QY 639 LGEIEAGANLL 649

Db 641 T-TIQAGASLM 650

RESULT 6

US-10-472-928-3912
; Sequence 3912, Application US/104722928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqMan99, version 1.03
; SEQ ID NO: 3912
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: PTS system, IABC components
; OTHER INFORMATION: Cellular location: membrane
; OTHER INFORMATION: Similar to strain R6 sequence 15903741 (0.E+01)
US-10-472-928-3912

Query Match 28.4%; Score 949; DB 17; Length 655;
Best Local Similarity 34.9%; Pred. No. 9.8e-73;
Matches 234; Conservative 128; Mismatches 253; Indels 56; Gaps 19;

QY 7 AQRILRDIGEDNIVAAAHCAITRLRLVLTCDTVDKROSLDDPDLKGFTEGMPQIIIVG 66
Db 8 AKDLLOAIGKENYAVVTHCATRMRFVLGDDKKANVKAIESIPAVKGFITNAGQFOVILIG 67
QY 67 PGDVHVFKELDATSKDIIVSTEQIKDVANNANWFSRAKVYIADIFVPLIPIIVGGGL 126
Db 68 -NDVPIFYNDP-TAVSGIEGVSKKAASAKSNQNVGVGWTTLTAIEFTPIIPALIVGL 125
QY 127 LMAINNVL--VAODLFGPQSLVEMFPQISGVAMINIMASAPFAFLPVLVGTATKRRFG 184
Db 126 ILGFNVLEGVWMSLDKTTESQFVAGVNHFLMLPGEALFQFLPVGITMSVSRKMG 185
QY 185 NEFLGAGIGMANVPFTLVNGVDVAATMTAGEMP--MMSLFGDLVDAQYQGTLPVLVVS 242
Db 186 SQTIGIVIGICLVSPQLNAAVASTPADIAANWVNGFYTVNRIGYQAOVIRPALLAG 245
QY 243 WILATIEKRLHRLMGTAFLITPVLTLLTGFLTFIAGPAMRWGDLAHGLOGLYDF 302
Db 246 LSLSTLEIFMWHKRIEIVISMIFVPLSLIPALILAHVTLGP---IGWTIGQGLSSVLA 301
QY 303 G--GPVG---GLFGLVSPVITVIGLHOSFPPIELELF--NOGGSFIPTASMANIAGAA 356
Db 302 GLTGPVKMLFGAIFGALVAPVITIGLHMTNADPOLADAGGTALMPMIALSNIAQSA 361
QY 357 CLAVFELAK--SEKGLAGAGSVAVLGTTEPAIFGVNLRRLRPYIGIGTAAGALIA 415
Db 362 VFAVYFMHDEREQAQVSLPATISAVLGTTEPALFGVNVKTYIPFVAGMTGALAGMLSV 421
QY 416 LEDIKAVNALGAAGFLGVNSIDAPDMWF---LYCAVVTFFVIAFGAALVGLYVRRNGS 471
Db 422 TTNVTAASIGIGLPGILISIQOYMLPFAGTMLVAIVPMLTLF-----PFRKAGL 472
QY 472 IDPDATAVPVAGTTKAAE---APAEFSNDSTI-----IOAPLTGEAIALSSVSDAM 521
Db 473 FTK-----TEGDTNLQAEFVAQBEAFVNHEPVELTSVEIISPLTQVKELSQATDPI 525
QY 522 PASGKLGSVAIVPTKQGLVSPVSGKIVVAPPSGHAFAVTRKAEQDSNVDILMHIGPDTV 581
Db 526 PASGVMGGGLVIEPSQGLTSPVNGTVVLPFTKHAIGI--VSDRG--VELLIHIGMDIV 581

```

Db      111 IPAMGGGKMLKVYLII-----PMLGILGSDGTTAFLTFPGDAPYFPLLLAYS 161
Qy      178 ATKAFGGNEFLGAGTIGMAWVPTLVNGYDVAAWTAGEMWMSLFGILDVAQAQGYQTVLP 237
      162 ASQGLKXTSTLAMSVAQVLLHPNFVQ-----MVQSGNPL-SLFGVPVLPASYGSSVVP 213
Qy      238 VLAVSWLLATIEKFLHKLMLGTADFLITPVYLLTLGLFTLFAIGPMAWVGDLLAHGLQ 297
      214 IIMVWMLKYEIKMAIKLTPAVTKSPLOPFLTVLLVSSCIALVAVGPGVIVGEGLSNLVG 273
Qy      298 GLYDFGPGVGLFELGVYSPVITVGLHQSPPIELBLFNQGSPIFATASMA----- 349
      274 QMYDVAMLTALIGALMPITVMTGMHMAP-----IFLAASITAPDVLIILP 321
Qy      350 -----NIAQGAACLAIVEFLAKSEKTLKGLAASGVSAVL-GITEPAIFGVMLRLMPFYIG 403
      322 AMLGSNLAQGAASNAVALSKRNNTKOIAFAAGSALMLAITEPALYGLVTKKPLY-- 379
Qy      404 IGTAAIGGALLALF---DIKAVAGAAGLV---VSIDAP-DWVMEFLCAVVTPIYAF 455
      380 --AMIGGAGLAGFAGLTSVKAYLEAVPSIALPQFIYSPVPSNIVALIAVIAVITF 437
Qy      456 GAALAYGLVYRRNGSIDPDATAAPVAGTKAEAPAFPSNDSTIOAPDLGEMALAS 515
      438 VLAVIFG-----IDEESS-----SNLEKAVSVNKKMIF-SPIGSEIILPLS 478
Qy      516 SVSAPMPSAGKLGSGVAIVPTKGQVSVAGKIVVAPPSGHAAPVTKAEDGSNVDLHM 575
      479 DVQKRTSDKLGIGGVAIIPSEKVVAPPROKTTNIPFTGAIGL--KSPBG--VELLIH 534
Qy      576 IGFDTVNLNGTHFNPLKKQGEVYKAGELCEPDIDAIIKAAGYEVTPIVSN 627
      535 IGLDTVELKQGGFISHVEEGDRVFNQGLIFEMDLNLIKTKGYETVPVIVTN 586

RESULT 8
US-10-474-776-253
; Sequence 253, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLY-
; TITLE OF INVENTION: ANTIGENS AND USES THEREOF
; FILE REFERENCE: AM100649-PT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 253
; LENGTH: 612
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-253

Query Match      23.9%; Score 799.5; DB 16; Length 612;
Best Local Similarity 31.4%; Pred. No. 7e-60;
Matches 205; Conservative 119; Mismatches 237; Indels 91; Gaps 18;

Qy      1 MDHKDLAQRILRDIIGEDNIVAAAHCASTRLLVLYKDTKYDVDRGSLDDDPKGTFFEGCM 60
      1 MSYKDTQKILDIVIGGEKNVNRVTHCVTRLRLLEIKCENLVNDDVYKKIPGVIGIMKKNQ 60
Qy      61 FQIIIVGSDVDHYEKE---LDDATSKDIAVSTEQLKDVANNANWPSRAVKVLAIDIFVPL 117
      61 YQIIILG-NDVANYKEVVKLGNFESDSV-----VOGKKNIILRIIEIYIAGSNTP 110
Db      118 IPIVVGGLMAINNVLAODLFGPOSIVEMFPOISGVAMINIMASAPAPLPLVIGFT 177
      111 IPAMGGGKMLKVYLII-----PMLGILGSDGTTAFLTFPGDAPYFPLLLAYS 161
Qy      178 ATKAFGGNEFLGAGIGMAWVPTLVNGYDVAAWTAGEMWMSLFGILDVAQAQGYQTVLP 237
      162 ASQGLKXTSTLAMSVAQVLLHPNFVQ-----MVQSGNPL-SLFGAPVLPASYGSSVVP 213
Qy      178 ATKAFGGNEFLGAGIGMAWVPTLVNGYDVAAWTAGEMWMSLFGILDVAQAQGYQTVLP 237
      162 ASQGLKXTSTLAMSVAQVLLHPNFVQ-----MVQSGNPL-SLFGAPVLPASYGSSVVP 213

```



```

; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74439
; LENGTH: 620
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74439

```

```

Query Match      23.5%; Score 785.5; DB 15; Length 620;
Best Local Similarity 30.7%; Pred. No. 1.1e-58;
Matches 195; Conservative 127; Mismatches 263; Indels 51; Gaps 13;

```

```

QY 1 MDKDLAQRILRDIGEDENIVAAACATRLRLVLTQTKVDNRQSLDDDDDLKQTFETG4 60
DB 1 MTYQETAKAILAANGSKTNIQRYTHCVTLRLVLRKDEKXQDQVAISNIGWAKNQ 60
QY 61 FOIIVPGDGDH--VFKELDATSKDIAVSTEQLDKDVVANNAMFSAVKVLADIFVL 117
DB 61 YQIIIG-NDVNNVYQAFSLGHFDNDQDEHSSKAKSIL-----ERLIETIAGVITPL 112
QY 118 IPIIVGGILMAINNVLAQDLFGPSLYEMFPQISGVAMINLMASAPAPLPVLVGT 177
DB 113 IPIIVGGILMAINNVLAQDLFGPSLYEMFPQISGVAMINLMASAPAPLPVLVGT 177
QY 178 ATKRFEGNEFLGAGIGMAVPTLVNGYDVAATMTAGEMPMMSLFGDLVAQGYQGTVP 237
DB 164 AAKRFVTVTLAATIGILHFAV-----AMVAEGKPL-TLFGAPVTPASYSGSVIP 215
QY 238 VLVSMTIATIEKFLHRLMGTADFLITPVLTLTLGFLTLFIAIGPAMRWGDLAAGLQ 297
DB 216 ILMVNYLMQYIEKVVRLVPSWKSFLQPTLILISGLALVVGGLVGIIGGGSNTML 275
QY 298 GLYDGGPRVGLFLGLVYPIYITGTHQSPPIELEFNQGGSFITATMSA-NIAQGA 356
DB 276 ALYHVAWPLALAILGLIMPLVMTGMHMAFAPFLASVATDVLILPMLSLNLAQGA 335
QY 357 CLAVEFLAKSEKLGKLAGASGVSAVL-GITEPAIFGVNLRMPFYTIGTALAIGGALA 415
DB 336 SLAAVAFKTKQKOTROVALAAGISALLAGITEPLVYTLKFKPLAANISGGLVGAFLG 395
QY 416 LFDIKAVALLGAAGFLGVNSIDAP---DWVMPLVCAVTFVIAFGAALVGYLYVRNGS 471
DB 396 FVNINASYTFVPSIIGLPQYINPSGGANFTNLIACTATIVLAFSLTWFMG----- 446
QY 472 IDPDATAPAVPAGTTKAEKAPAFBSNDSTIIQAPLTGAILASSVSDMFPASGKLGCV 531
DB 447 IDEE---SEKQVSVADMSQVKSGLSTKOT-LYAPWTGEMLFISEVPDETFSKLGEGF 502
QY 532 AIVPTKGVSPVSGKIVVAFPSGHAFAVRTKAEDSNVDIIMHIGFDVNNNGTHFNPL 591
DB 503 AILPSEGEVYAFPDGEVITFFPTKAAVALK---NTRGVEVLIHGIDIVELKGGQFEOL 558
QY 592 KKQGDVAKAGELLCEFDIDAIRAAGEVTTPIVNSN 627
DB 559 VSVGDVVKRQALMKMDIDIFITSKGLISPIVVVN 594

```

```

RESULT 11
US-10-282-122A-70209
; Sequence 70209, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykend, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.

```

```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70209
; LENGTH: 379
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-10-282-122A-70209

```

```

Query Match      22.1%; Score 738; DB 15; Length 379;
Best Local Similarity 39.5%; Pred. No. 6.9e-55;
Matches 144; Conservative 89; Mismatches 116; Indels 16; Gaps 5;

```

```

QY 97 ANNANFSAVAVTLADIFVPLIPIVGGGLMAINNVLAQDLFGPSLYEMFPQISGVA 156
DB 1 AOKGNPVRFLITLGDIFPIPIPAIVTAGLLMGINNLTLMKGLFPGKALIEVYPOIADIS 60
QY 157 EMINMASAPAFELVUVGFTATKRFEGNEFLGAGIGAMAVPTLVNGYDVAATMTAGEM 216
DB 61 NIIINIASAFTEFLPALIGWSSRVFGSPPIIGIVGLIMHPOLVSQYDLA---KQNI 116
QY 217 PMSLFGDLVDAQYQGTVPVLVVSMTIATIEKFLHRLMGTADFLITPVLTLTLTGPL 276
DB 117 PWNLFGLIEKLANQOGQVLPVLIAYLAKIEKGLNKNVHDSIGMLVGPVALLVTGFL 176
QY 277 TFIAGPAMRWVGDLLANGLGGLYDFGPGVGLFLGLVYSPVITGTHQSPPIELELFN 336
DB 177 AFIIIGPYALLIGTIGTSGVTFFIFQHAQMLGAIVGLYAPLVITGLHMFPLAVFQGLMG 236
QY 337 Q--GGSFIFATMSANINQGAACLAFFLASEKL---KGLAGASGVSAVGTETPAIFG 391
DB 237 SLGGTYLMPVIAISNICQSSAFAFAMFYKRRKRVKKEGLLTCISGMLGVTBPMWG 296
QY 392 VNLRLRMPFYIGIGAAIGALIALFDIKAVALLGAAGFLGV--VSIDAPDMWFLVCAY 448
DB 297 VNLPLKFFIATLSTCVLGAIVGMNV---LGRVGVGVAFATISIQKEFFPVYLIARA 352

```

```

RESULT 12
US-09-738-626-6961
; Sequence 6961, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI

```


Qy 615 AGCEVTTPIVSNYKKTGPNTYGLGEIRAGANILNVAK 654
Db 642 KDLPLITPVVSNNAKFGIEBGPADQANSTTVIVNKG 681

RESULT 14
US-10-282-122A-57628
Sequence 57628, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 57628
LENGTH: 451
TYPE: PRT
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (223)..(223)
OTHER INFORMATION: X=any amino acid
US-10-282-122A-57628

Query Match 18.8%; Score 627; DB 15; Length 451;
Best Local Similarity 33.4%; Pred. No. 3,4e-45;
Matches 158; Conservative 82; Mismatches 199; Indels 34; Gaps 12;

Qy 1 MDHDLAQRILRDIGEDENIVAAHCATRLRLVLTOKTKVDROSLDDPDLKGTETGGM 60
Db 1 MEYKELGKOTILANVGSESVSLAHCATRLR--FDDSKADSAIRNLKGVGVNNGSQ 58
Qy 61 FOIIVGPDVDAVFELEDATSKDIASVTEQLKDVVANNAMFSRAVKVLADI FVPLPI 120
Db 59 FOVIIG-GVYKFEVFAEIDTL--GDFSASKESSSE--ANNKGIVSKVLDITIGIFVPIVPA 113
Qy 121 LVGGGLMANNVLAQDLFGPSLVEMFPQISGVAEMINLMAAPFAFLPVLYGFTATK 180

Db 114 LTGAGMLKALLALVLMEMTPESQYOF-----LNFIDGAFFYLPVLLASSAAK 164
Qy 181 RRGNEPFGAGIGAMAVPEPLVNGDVAAITMAGEMPMSLFGDLVAQAGYGTLPVLV 240
Db 165 KFGCNQYLAITLIGALLHPFT--AMIAARETGE--SLALFGLVPTLANYSVPIPL 220
Qy 241 VSMILATEKFLHKKMGTADELTPVLTLLTGFLFETAIAGPARWVDLLAHGLOGLY 300
Db 221 AIXFWSYIEPIIDKMPANAVRFMTPLTLTLVAITLIVIGF---LGTFLGNLGLSL 276
Qy 301 DFGSPVGG---LFLGLVSPYIVITGLHQSPPIELDFNOGSGFIATASN--ANIAG 354
Db 277 EFVNTYASWLVPLVGAFTPLVMTGMHYGLIPGINMLATSGPDTVAGPMWGSNIAG 336
Qy 355 AACLAFFLAKSEKIKGAGASGVAUGITEPAIFGVNLRMRPFYIGICTAAGALI 414
Db 337 GAALAVAFRTKQLSVOKLAVSSGLSAFNGITEPALYGINMRKKPLIASWGGAAGLEI 396
Qy 415 ALFDIKAVALGAAGFLGVVS-IDAPDMVFLVCAV---VTFVIAAGAAIAYGL 463
Db 397 GIMGVRYAQAVALPILALPSFFGKGLSNFYAALGCAIAFVGVFIASFTLGI 449

RESULT 15
US-10-474-776-367
Sequence 367, Application US/10474776
Publication No. US20040110181A1
GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYP
FILE REFERENCE: AM100649-PCT
CURRENT APPLICATION NUMBER: US/10/474,776
CURRENT FILING DATE: 2003-10-14
NUMBER OF SEQ ID NOS: 752
SOFTWARE: PatentIn version 3.1
SEQ ID NO 367
LENGTH: 381
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-474-776-367

Query Match 17.5%; Score 584; DB 16; Length 381;
Best Local Similarity 34.6%; Pred. No. 1.3e-41;
Matches 139; Conservative 81; Mismatches 146; Indels 36; Gaps 9;

Qy 263 LITPVLTLTGFLFFAIGPAMRWGDLAHGLOGLVDFGPGVGLFGLVYSPVITG 322
Db 1 MVTPTVLNWSILGLFVIGPVFHVENVYIILATAIISMPPGLGFLIGVHQLIVSG 60
Qy 323 LHQSPPIELDFNOGSGFIF-ATASMANIAGAACLAFFLAKSEKIKGAGASVAS 381
Db 61 VHHIFNLLEVQLADHDANPFAAITAAMTAQGAATVAVGXTKPKLTIAFPALSAF 120
Qy 382 LGITEPAIFGVNLRMRPFYIGICTAAGALIALFDIKAVALGAAGFLGVVISIAPMV 441
Db 121 LGITEPAIFGVNLRMRKPFPLSLIAGIAGGLASI-----LGLAGTNGNGTITIPGTM 173
Qy 442 -----MFLVCAVTFVIAFGAIAVGLVYVRNGSIDPATAAPVAGTTKAEAE-A 492
Db 174 YVGNQOLQOYLIMAVVSALGFALTMYMGY-----EDVDATYAAKRAVEKEBEVA 226
Qy 493 PAEFNSDSTIIQAPLTJGALIASVSDAMFASGKLGSGVAIVPTKQLVSPSGKIIVAF 552
Db 227 PAALQNETLV--TPIGVVVALADVNDPVFSSGAMGQGIIVKPSGVVYAPADAEVIAF 284
Qy 553 PSGAFAFVTRKAEQSDNDILMHIGFPTVYNANGTHFNPLKXQGBVKKAGELCEPDIAI 612
Db 285 PTGHAFGKTR---NGAEVLIHVGIDIVSNMGDGFETKVAQGNKVKAGADVLTGTDNSKI 340
Qy 613 KAAGEVTTPIVSN---YKKTGPNTYGLGEIRAGANILNV 651
Db 341 AAAGLDITTWIVTNTGDIYASAVAVAT---GSVAKGDAVIEV 379

Fri Mar 11 12:40:57 2005

us-10-019-284b-2.rapb

Page 10

Search completed: March 7, 2005, 22:08:07
Job time : 139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 22:05:18 ; Search time 44 Seconds

(without alignments)
1445.438 Million cell updates/sec

Title: US-10-019-284B-2

Perfect score: 3342

Sequence: 1 MDHKLQRIIRDIGEDNI.....IAGANLNVAKKAVPATP 661

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1609.5	48.2	651	2 S44257	phosphotransferase
2	1478.5	44.2	664	1 B32243	phosphotransferase
3	1227	36.7	480	2 S39978	scra protein - Sta
4	1178	35.2	480	2 D90038	PTS system, sucrose
5	1098.5	32.9	627	2 H96951	fusion, PTS system
6	1093	32.7	479	2 F82432	PTS system, sucrose
7	1069	32.0	479	2 JQ0781	sucrose uptake pro
8	1013.5	30.3	470	2 H83926	PTS system, trehal
9	999	29.9	632	2 C69725	phosphotransferase
10	990.5	29.6	470	2 S68599	phosphotransferase
11	960	28.7	627	2 P95200	PTS system IIABC C
12	957	28.6	617	2 AB1167	phosphotransferase
13	956	28.6	627	2 E98067	phosphotransferase
14	949	28.4	655	2 C95220	trehalose PTS syst
15	934	27.9	705	2 A99084	phosphotransferase
16	933.5	27.8	475	2 C89813	hypothetical prote
17	930.5	27.8	630	2 H83686	PTS system, beta-g
18	929.5	27.8	636	2 C83724	PTS system, beta-g
19	927	27.7	633	2 AC1436	PTS system, beta-g
20	924.5	27.7	634	2 AD1078	PTS system, beta-g
21	901	27.0	609	2 I40406	beta-glucoside per
22	896	26.8	631	2 B42603	beta-glucoside-spe
23	887	26.5	609	2 T47097	hypothetical prote
24	860.5	25.7	628	2 D97073	PTS system, beta-g
25	859	25.7	617	2 AC1421	beta-glucoside-spe
26	829	24.8	618	2 AC1204	phosphotransferase
27	824	24.7	625	2 C25977	phosphotransferase
28	822.5	24.6	636	2 D86807	hypothetical prote
29	816	24.4	456	2 S62331	phosphotransferase

30	801.5	24.0	612	2 A97935	hypothetical prote
31	799.5	23.9	612	2 B95067	hypothetical prote
32	781	23.4	458	2 H83881	PTS system, sucrose
33	770.5	23.1	455	1 WQ8BST	phosphotransferase
34	751.5	22.5	460	2 A39938	phosphotransferase
35	747.5	22.4	640	2 AB1423	beta-glucoside-spe
36	719	21.5	494	2 AG1231	PTS system trehalo
37	706	21.1	459	2 J00293	levansucrase synth
38	703.5	21.1	494	2 AF1585	PTS system trehalo
39	669.5	20.0	483	2 A10449	protein-Npi-phosph
40	659	19.7	473	2 A98281	trehalose specific
41	655	19.6	473	2 C65236	phosphotransferase
42	652	19.5	473	2 A86122	PTS system enzyme
43	641	19.2	478	2 B82263	PTS system, trehal
44	630.5	18.9	372	2 I39868	sac operon regulat
45	620	18.6	472	2 AF1058	protein-Npi-phosph

ALIGNMENTS

RESULT 1	S44257	phosphotransferase system enzyme II (EC 2.7.1.69), sucrose specific - <i>Pedococcus pentosaceus</i>
C/Species:	<i>Pedococcus pentosaceus</i>	
C/Date:	19-Mar-1997	#sequence_revision 19-Mar-1997 #ext_change 09-Jul-2004
C/Accession:	S44257	
R/Leenhouts, K.K.J.; Bolhuis, A.A.; Kok, J.J.; Venema, G.G.		
A/Description:	The sucrose and raffinose operons of <i>Pedococcus pentosaceus</i> PP1.0.	
A/Reference number:	S44252	
A/Accession:	S44257	
A/Molecule type:	DNA	
A/Residues:	1-651 <LEE>	
A/Cross-references:	UNIPROT:P43470; EMBL:Z32771; NID:g493728; PIDN:CAA83668.1; PID:g475;	
A/Genes:	scra	
C/Superfamily:	phosphotransferase system enzyme II sucrose-specific; phosphotransferase	
C/Keywords:	phosphotransferase	
F/488-640/Domain:	phosphotransferase system glucose-specific enzyme II, factor III homo	
Query Match	48.2% Score 1609.5; DB 2; Length 651;	
Best Local Similarity	49.6%; Pred. No. 1.9e-99;	
Matches 331; Conservative 114; Mismatches 196; Indels 27; Gaps 8;		
QY	1 MDHKLQRIIRDIGEDNIYAAACATRLRLVLTQDVDRQSLDDPDLKGPETGM 60	
DB	1 MHQGVADRVNLA-I-GKNNIQAAHCATRLRLVITDESKIDQALDDADVAGTFETNGQ 59	
QY	61 FOITVPGDVDFEYKELDDATSKDIAVSTEQLKDVVA--NNANFSAVKVLTADIVPLI 118	
DB	60 YOIIIGPGVDVCKYDALIVKGLK-EVTPDDDKAAVAAGQNKPNLMDFLKVISDFIPV 118	
QY	119 PLTVGGGLMANNVTVNODLFGPOSIVEMFQISGVEMIMLASAPAPLPVTVGFTA 178	
DB	119 PALVAGGLMALNNVLTAEHLFMASVVEVYGLGIMINAMASAPFTPIILGFSFA 178	
QY	179 TRKFGNEFLGAGIGMAWFPPLVANGYDVAATMTAGEMPMSLFGLDVQAQGYQVTLRV 238	
DB	179 TRKFGNPTLGTATGMIWVLPVLVNGYSVATLMAAGKVMYVNFGLHVAQAGYQGVQLV 238	
QY	239 LVVSMILATIEKFLHRLMGTDPLITPVTLTLGLFLTAIGPAMRWGLLAHGLQG 298	
DB	239 LGVAVILATIEKFPFKHKGAPDFTTPFAIVITGFLFTIVGVLKRVSPALNTGLVG 298	
QY	299 LVDFGPGVGLLFGVVSPIVITGHOSFPPIELELF-----NQGSFIPATASMANIAG 354	
DB	299 LVNSTGVMGTFGLGLSAIVITGLHQTTPALFETQLANVATGSGSIFPVAASANIQG 358	
QY	355 AACLAFFLAKSEKLKGLAGASVSAVLGITPEPAIFGVNLRWPFYIGITGAAGALI 414	
DB	359 AATLAIFFATKSKQKALTSAGVSLGITEPALFGVNLKKPFVFPAIASGASAVL 418	

A:Cross-references: UNIPROT:O9KLT8; GB:AE004395; GB:AE003853; NID:g9658068; PIDN:AAF9655
A:Experimental source: serogroup O1, strain N1661; biotype E1 tor
C:Genetics:
A:Gene: VCA0653
A:Map position: 2
C:Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

Query Match 32.7%; Score 1093; DB 2; Length 479;
Best Local Similarity 47.8%; Pred. No. 3.1e-65;
Matches 222; Conservative 85; Mismatches 151; Indels 6; Gaps 3;

```
OY 1 MDHKLADRIIDIGEDNIVAAACATRLRLVLTAKDTQDVDRQSDDDPDLKGTFTGGM 60
DB 1 MDYPIAKQLESLSGKSNITQALACATRLRLVLDNDERQINSALIESIQGVKGQKVAQG 60
OY 61 FQIIVGPDVDHVFELDDATSKDIAVSTQKDVANNANMFSAVYLVADIFVPLPI 120
DB 61 FQIIFGSGIVNQVAVEMAKLTGL-VEMSTNDVASAGAEKQNAQRAVYKGLSDIFVRIIPA 119
OY 121 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVGFTAT 179
DB 121 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVGFTAT 179
OY 120 IVAGGLMGLFVNLTPAGLFTGQSLIDANPGLADLASINTEFANAPFVYLPVLLAFSAS 179
DB 120 IVAGGLMGLFVNLTPAGLFTGQSLIDANPGLADLASINTEFANAPFVYLPVLLAFSAS 179
OY 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
DB 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
OY 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
DB 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
OY 240 VWSMILATIEKFLHRLMGTADFLITPVLTLTLGFTFLFIAIGPAMRWGDLAHGLOGL 299
DB 240 VSAIYIAKIKENGLRKTIVPSVDNLTLPMLAIPITGFLFTVVGFLTRDVGFLDALNWL 299
OY 300 YDFGPGVGLLFGLVSPVITGLHQSFPPIELEF---NQGSPFIPATASMANIAGA 355
DB 300 YDFGPGVGLLFGLVSPVITGLHQSFPPIELEF---NQGSPFIPATASMANIAGA 355
OY 300 YDSAPFVGAGLFGFIVAPVITGMHSFIAETQLADIVTTGGFPIPIAMSNVAGA 359
DB 300 YDSAPFVGAGLFGFIVAPVITGMHSFIAETQLADIVTTGGFPIPIAMSNVAGA 359
OY 356 ACIAVFLAKSEKLGKLAGAGSVAVLGITPEPAIFGVNLRMRPFYIGITRAIGALIA 415
DB 356 ACIAVFLAKSEKLGKLAGAGSVAVLGITPEPAIFGVNLRMRPFYIGITRAIGALIA 415
OY 360 AALAVGWTKETKLGKGLAIPSGVTALLGITPEPAMGVNLRKYPFIATACGAAIASAFT 419
DB 360 AALAVGWTKETKLGKGLAIPSGVTALLGITPEPAMGVNLRKYPFIATACGAAIASAFT 419
OY 416 LFDIKAVALLGAAGFLGVVSIDAPDMWMLVCAVVTFFIARGAIAI 459
DB 420 LFNVAQALGAAGLPGIISINPQIGYIMGMAISFVAAFALT 463
```

RESULT 7

Query Match 32.0%; Score 1069; DB 2; Length 479;
Best Local Similarity 46.1%; Pred. No. 1.2e-63;
Matches 216; Conservative 94; Mismatches 151; Indels 8; Gaps 5;

Query Match 32.0%; Score 1069; DB 2; Length 479;
Best Local Similarity 46.1%; Pred. No. 1.2e-63;
Matches 216; Conservative 94; Mismatches 151; Indels 8; Gaps 5;

Query Match 32.0%; Score 1069; DB 2; Length 479;
Best Local Similarity 46.1%; Pred. No. 1.2e-63;
Matches 216; Conservative 94; Mismatches 151; Indels 8; Gaps 5;

```
OY 1 MDHKLADRIIDIGEDNIVAAACATRLRLVLTAKDTQDVDRQSDDDPDLKGTFTGGM 60
DB 1 MDYPIAKQLESLSGKSNITQALACATRLRLVLDNDERQINSALIESIQGVKGQKVAQG 60
OY 61 FQIIVGPDVDHVFELDDATSKDIAVSTQKDVANNANMFSAVYLVADIFVPLPI 120
DB 61 FQIIFGSGIVNQVAVEMAKLTGL-VEMSTNDVASAGAEKQNAQRAVYKGLSDIFVRIIPA 119
```

```
DB 61 FQIIFGSGIVNQVAVEMAKLTGL-VEMSTNDVASAGAEKQNAQRAVYKGLSDIFVRIIPA 119
```

```
OY 121 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVGFTAT 179
DB 121 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVGFTAT 179
OY 120 IVAGGLMGLIYNLLTAQGLFDGSLIEANPGLTDLANMINITFANAPFVYLPVLLAFSAS 179
DB 120 IVAGGLMGLIYNLLTAQGLFDGSLIEANPGLTDLANMINITFANAPFVYLPVLLAFSAS 179
OY 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
DB 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
OY 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
DB 180 KRFGNNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGTVPVL 239
OY 240 VWSMILATIEKFLHRLMGTADFLITPVLTLTLGFTFLFIAIGPAMRWGDLAHGLOGL 299
DB 240 VSAIYIAKIKENGLRKTIVPSVDNLTLPMLAIPITGFLFTVVGFLTRDVGFLDALNWL 299
OY 300 YDFGPGVGLLFGLVSPVITGLHQSFPPIELEF---NQGSPFIPATASMANIAGA 355
DB 300 YDFGPGVGLLFGLVSPVITGLHQSFPPIELEF---NQGSPFIPATASMANIAGA 355
OY 300 YDSAPFVGAGLFGFIVAPVITGMHSFIAETQLADIVTTGGFPIPIAMSNVAGA 359
DB 300 YDSAPFVGAGLFGFIVAPVITGMHSFIAETQLADIVTTGGFPIPIAMSNVAGA 359
OY 356 ACIAVFLAKSEKLGKLAGAGSVAVLGITPEPAIFGVNLRMRPFYIGITRAIGALIA 415
DB 356 ACIAVFLAKSEKLGKLAGAGSVAVLGITPEPAIFGVNLRMRPFYIGITRAIGALIA 415
OY 360 AALAVGWMSKQKMGKGLAIPSGVTALLGITPEPAMGVNLRKYPFIATACGAAIASAFT 419
DB 360 AALAVGWMSKQKMGKGLAIPSGVTALLGITPEPAMGVNLRKYPFIATACGAAIASAFT 419
OY 416 LFDIKAVALLGAAGFLGVVSIDAPDMWMLVCAVVTFFIARGAIAI 463
DB 420 MFWNAQALGAAGLPGIISIT-TPDKIGYIAGMVIATPLAVLTIVLGI 467
```

RESULT 8

Query Match 30.3%; Score 1013.5; DB 2; Length 470;
Best Local Similarity 44.4%; Pred. No. 5.8e-60;
Matches 204; Conservative 90; Mismatches 150; Indels 15; Gaps 4;

Query Match 30.3%; Score 1013.5; DB 2; Length 470;
Best Local Similarity 44.4%; Pred. No. 5.8e-60;
Matches 204; Conservative 90; Mismatches 150; Indels 15; Gaps 4;

Query Match 30.3%; Score 1013.5; DB 2; Length 470;
Best Local Similarity 44.4%; Pred. No. 5.8e-60;
Matches 204; Conservative 90; Mismatches 150; Indels 15; Gaps 4;

```
OY 3 HKDLAORILRDIGEDNIVAAACATRLRLVLTAKDTQDVDRQSDDDPDLKGTFTGGM 62
DB 4 YKKEVNAIVEAIGADNITQATHCVTRLRFLVLRDEGKVDQKLSLDIVKGSFSTNGQFQ 63
OY 63 IIVGPDVDHVFELDDATSKDIAVSTQKDVANNANMFSAVYLVADIFVPLPI 116
DB 64 VIKKEVNAIVEAIGADNITQATHCVTRLRFLVLRDEGKVDQKLSLDIVKGSFSTNGQFQ 63
OY 117 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVG 175
DB 117 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVG 175
OY 117 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVG 175
DB 117 LVGGGLMAINNVLVAODLF-GPQSLVEMFPQISGVAEMINIMASAPAFPLPVVG 175
OY 176 FTATKRFQNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGT 235
DB 176 FTATKRFQNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGT 235
OY 177 MSAYKRFQNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGT 236
DB 177 MSAYKRFQNEFLGAGIGAMVFPFLTVNGYDVAATWTAGEMPWMSLFGDVAQAQYQGT 236
OY 236 LPLVTVSWILATIEKFLHRLMGTADFLITPVLTLTLGFTFLFIAIGPAMRWGDLAHG 295
DB 237 LPLVTVSWILATIEKFLHRLMGTADFLITPVLTLTLGFTFLFIAIGPAMRWGDLAHG 295
OY 296 LQGLYDRGPGVGLLFGLVSPVITGLHQSFPPIELEFNG-GGSFIPATASMANIAG 354
DB 296 LQGLYDRGPGVGLLFGLVSPVITGLHQSFPPIELEFNG-GGSFIPATASMANIAG 354
```

Db 297 FTSIFAALVGVGLGYLTAFLVYGMHTFLAVALDQLIGTGFPLVLSIAG 356
Qy 355 AACLAVFLAKSEKLGAGASVAVLITTEPAIFGVNLRMPFYIGTAIGALI 414
Db 357 SAALAMFATPBKLGSLSSAVSLGITTEPAMGVNIRYFPVCAIISAIGARI 416
Qy 415 ALPDIKAVALLGAAGFVGVSIDAPDMVMLVCVAVTFVI 453
Db 417 TVNGVLANSIGVGGLPGRFISIQAGFVGVFIGVIAFIL 455

RESULT 9
C69725
phosphotransferase system enzyme II (EC 2.7.1.69) phosphoenolpyruvate-dependent, trehalo
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence, revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: C69725; S67929; J05037; I40497; S67864
R/Kunzt, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Ertlan, K.D.; Ertlan, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
Jech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulo, M.F.
Kestler, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maruda, S.; Mausel
Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Schlitch, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowski, A.; Serot
Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69725

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-470 <KUN>

A/Cross-references: UNIPROT:P39794; GB:299108; GB:AL009126; NID:G2633055; PIDN:CAB12609.

A/Experimental source: strain 168

R/Helfert, C.; Gotsche, S.; Dahl, M.K.
Mol. Microbiol. 16, 111-120, 1995

A/Title: Cleavage of trehalose-phosphate in Bacillus subtilis is catalysed by a phospho-
A/Reference number: I40497; MUID:95379486; PMID:7651129

A/Accession: S67929

A/Molecule type: DNA

A/Residues: 324-362, 'L', 364-464, 'G', 466-470 <HEL>

A/Cross-references: EMBL:X80203; NID:G580941

R/Schoeck, F.; Dahl, M.K.
Gene 175, 59-63, 1996

A/Title: Analysis of DNA flanking the trea gene of Bacillus subtilis reveals genes encod
A/Reference number: J05037; MUID:97074649; PMID:8917076

A/Accession: J05037

A/Molecule type: DNA

A/Residues: 1-139, 'S', 141-362, 'L', 364-464, 'G', 466-470 <SC2>

A/Cross-references: EMBL:254245; NID:G1000450; PIDN:CA91014.1, PID:G1000451

C/Comment: This enzyme functions as the specific trehalose transporter. It belongs to th
C/Genetics:
A/Genes: trep, treB
C/Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II
C/Keywords: phosphotransferase, sugar transport system

F/113-133/Domain: transmembrane #status predicted <TM1>
F/160-181/Domain: transmembrane #status predicted <TM2>
F/184-203/Domain: transmembrane #status predicted <TM3>
F/230-245/Domain: transmembrane #status predicted <TM4>
F/263-286/Domain: transmembrane #status predicted <TM5>
F/305-325/Domain: transmembrane #status predicted <TM6>
F/375-399/Domain: transmembrane #status predicted <TM7>
F/402-421/Domain: transmembrane #status predicted <TM8>

Query Match 29.9%; Score 999; DB 2; Length 470;
Best Local Similarity 43.9%; Pred. No. 5-4e-59;
Matches 203; Conservative 93; Mismatches 160; Indels 6; Gaps 4;

Qy 7 AQRIRLDIGGEENIYAAAHCAIRLRLVLTQTDVDRDSDDDPDLKGTFTGTGMPIY 66
Db 8 AAOIEAVAGENIAAACHCTVTRFLRFLIDSKVQOEMDQIDVVGKSPSTGQFQVVG 67
Qy 67 PGDVHVFPELDDASKDIASVT-BOLKDVANNANWSPRAVKVADIFVPLIPILYVG 125
Db 68 QCTVAKVIAELVKETG--IGESTKDVKKASERKNAPLORAVKTLADIFILPLAVTG 125

Qy 126 LIMAINNVLVAODL-FGPOSLVEMPOISGVLEMNINLMAAPAFPLVVGFTATRF 184
Db 126 LIMGINNLTLAGIFSTRTSIQVYVQWADLANMINLTAFTPLPALIGMSAVRFG 185

Qy 185 NEFLNAGIGMANVFTLVNGVDVAATMTAGEMPMSTGLDVAQYQCTVPLVVSPI 244
Db 186 NPLDGIIVGVMLVHDLNLNMGVGAABOSGEIPVNNLFGLEVQYQGVPLILASVM 245

Qy 245 LATIEKPLHKLGMADFLITPVLTLTLTGFTLTAIGPAMRWGDDLAHGLQGLYDF 304
Db 246 LAKIEVPLKTRPEGIQLVAPITLLTGFPASFIITGFTPAIGVLTSGLSVFGSA 305

Qy 305 PVGGLLFGIVSPVITGTHOSFPPIELFNO--GGSFIFATASMANIQAACIAPVF 362
Db 306 ALGGLLGGFYSALVITGMHTFLAVDLQLIGSKLGGTFMPLALNSIAQSSALAMVF 365

Qy 363 LAKSEKLGAGASVAVLITTEPAIFGVNLRMPFYIGTAIGALIAFDIKAV 422
Db 366 IYKDEKQKGLSTLGSISAVLITTEPAIFGVNLRMPFYIGTAIGALIAFDIKAV 425

Qy 423 ALGAGFGLGVSIDAPDMVMLVCVAVTFVIAFGAIVGLY 464
Db 426 SVGVGVPEIRTSIMSQWGAFAIGMAIVLIVFAGTAVARF 467

RESULT 10
S68599
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Streptococcus sobr
N/Alternate names: sucrose-specific enzyme II
C/Species: Streptococcus sobrinus
A/Variety: strain 6715
C/Date: 14-Feb-1997 #sequence, revision 13-Mar-1997 #text_change 07-May-1999

R/Chen, Y.Y.M.; Lee, L.N.; LeBlanc, D.J.
Infect. Immun. 61, 2602-2610, 1993

A/Title: Sequence analysis of scra and scrb from Streptococcus sobrinus 6715.

A/Reference number: S68598; MUID:93273516; PMID:8500898

A/Accession: S68599

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-632 <CH>

A/Cross-references: EMBL:L06791

C/Genetics:
A/Genes: scra
C/Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
C/Keywords: phosphotransferase, sugar transport system

F/480-632/Domain: phosphotransferase system glucose-specific enzyme II, factor III homol

Query Match 29.6%; Score 990.5; DB 2; Length 632;
Best Local Similarity 35.5%; Pred. No. 2.8e-58;
Matches 236; Conservative 114; Mismatches 270; Indels 45; Gaps 12;

Qy 1 MDHKLQRIILRDIGEDNIVAAAHCAIRLRLVLTQTDVDRDSDDDPDLKGTFTGTGM 60
Db 1 MDNKOIAKEVIEALGGRDNRVSVAHCAIRLRLVMTDEAKIKEREMENDKYGAFFNSGQ 60

Qy 61 FOITVPGDVHVFPELDD--ATSKDIASVTBOUKOVVANNANWSPRAVKVADIFVPL 117
Db 61 YQITFGTGVNKTIDEVVDLGLPTS----STGEKQGEAAGQNMFGQMSRTFGDVFPV 115

Qy 118 IPIIVGGILMAINNVLVAODLFGPOSLVEMFQISGVLEMNIN-----LMSAPPAF 169
Db 116 IPIVATATGLFMGRLRLTNDPFLG-----FPGASKQIINNFIYQVLTDTATAF 166

[illegible]

RESULT 11
 P95200
 PHS system IIABC components [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C:Accession: P95200
 R:Retelain, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Unayam, L.A.; White, O.; Salberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfe,
 neon, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison
 A>Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: P95200
 A>Status: preliminary
 A:Molecule type: DNA
 A:Reads: 1-627 <KUR>
 A:Cross-references: UNIPROT:Q97PB8; GB:AE005672; PIDD:AAK75799.1; PID:g14973217; GSPDB:C
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP1722
 C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Best Local Similarity 33.3%; Pred. No. 3e-56;
Matches 222; Conservative 141; Mismatches 246; Indels 58; Gaps 14;

```
Oy      1 MHKDLAQRILRDIGGEDIIVAAACATRLRVLKPTQDVGRQSDDDDPDLKGFEETGM 60  
         |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Db      1 MNNGSLAKVLDALGERENVNSVAHCAITLRWVDDEEKINEVENLEKQGAFENSG 60  
  
Oy     61 FOIIVGPGEVDHVFKE---LDDATSKDIAVSTEQLDKVANNANMFSRAVKYLLADI FVPL 117  
         |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Db     61 YOIIFGTGVNKKRYEDVVVLGLPTS-----SKDDMKAEFAKAGGNFGQALIRTFGGVFPDI 115  
  
Oy    118 IDLVGGGLIMAINNVLAQDLFGPSQLVEVPFOISGAEMINMLASAPFAPFLPVLYGFT 177
```

```

Db      116  I P V I A T L E N G V R G L E N F A L E M ----- P L P E D F A T Y T O I L D T A F I I L P G L V W S 165
Oy      178  A T K R F G N E F L G A G I G M A M F P T L V N G V D A A T M T A G E N P M S L F G L D V A O A G Y O G T L P 237
Db      166  T F R V G G N P A V I G I V L G M M L V S G S L P N A M A V A --- O G G E V T A N M F G F - I P V V G I O G S V L P 221
Oy      238  V L V S M I I A T T E K F L H K R L M T A D F L I T P V L T I L L T G F L F I A I G P M R V S G L L A H G L Q 297
Db      222  A F I I G V A K E K A V R K V P D V I D L L T P F V L L W S I L G L E V I G P V H V E N T I I A T K 281
Oy      298  G L Y D G G P V G G L L F G I V S P I V T G H O S F P P I E L E F N O G S F I P - A T A S M A N I A O G A A 356
Db      282  A I L S P F E L G G L I G G V H O L I V S G V H I F E N L E V O L L A D H A N P F A I I T A A M T A G A A 341
Oy      357  C I A V F L A K S E K L G I A G A S V A V I G T E P A I F G V N L R M F P Y I G I G A I L G A L I A L 416
Db      342  T V A V G V K T N P K L K T L A P P A L S A F L G I T E P A I F G N L R R K P F F S I L G A I G G G L A S I 401
Oy      417  F D I K A V A L G A A G E F L G V S I D A P D M V ----- M P I V C A V T F V I A F G A I A A Y G L Y V R 468
Db      402  ----- I S L A T G M G I T I I P G M L Y V G N G O L P Q Y I L M A V S F A L G A L Y M F G Y ----- 449
Oy      469  N G S I D P D A T A A V P A G T T K A E A E - A P A E F S N D S T I I Q A P I T G A I A L S S V D A M F A S G K L 527
Db      450  -- E D E V D A T A A K R A E V A E K E E V A P A A L O N E T L V -- T P I V G V V A L A D V D P V F S S G A M 505
Oy      528  G S G V A I V P R K G O L V S P V S G K I V A P R P S G H A F A V R T A E D S A N D I L M H I G D F T N L N G T H 587
Db      506  G G G I V A K S O G V V A P A D A E V S I A P T G H A F G L K T R --- N G A E V L I H V G I D I V S M N O D G 561
Oy      588  F N P L K K O G D E V A G E L L C E F D I D A I K A A G Y E V T T P I V S N - - Y K K T G P A N T Y G L G E I E A 644
Db      562  F E T K V A O G K V A G A D V L G T F D S N K I A A G L D T T W I V T N T G D Y A S V A P A T --- G S V A K 618
Oy      645  G A N L I A N V 651
Db      619  G P A V I E V 625

```

RESULT 12
AB1167
phosphotransferase system (PTS) beta-glucoside-specific enzyme IIABc component homolog 1
C|Species: *Listeria monocytogenes*
C|Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C|Accession: AB1167
R|Glaeuer, P., Frangoul, L., Buchrieser, C., Amend, A., Baquero, F., Berche, P., Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauget, O.; Entian, K.D.; Fshni, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A|Authors: Kreif, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma
A|C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.|Title: Comparative genomics of *Listeria* species.
A|Reference number: AB1077; NUID:21537279; PMID:11679669
A|Accession: AB1167
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-617 <GLA>
A|Cross-references: UNIPROT:Q8Y904; GB:NC_003210; PIDN:GAC98816.1; PID:G16410127; GSPDB:
A|Experimental source: strain EGD-e
C|Genetics: none

A Jensen JMO0738
C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match 28.6%; Score 957; DB 2; Length 617;
Best Local Similarity 35.1%; Pred. No. 4.6e-56;
Matches 225; Conservative 124; Mismatches 228; Indels 64; Gaps 17;

QY 1 MDHKDLAQRLLRDIGGEDNIVAAAHCAITRLRLVLTQDKVDNRÖSLDDPDLKGFETGGM 60
 ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|
Db 1 MDYOKLAKELLAVNGEENRVSVHCATRLRFKLVNNEKADKKOIESISGVISVENNQ 60

QY 61 FQLTVG--PEDVDHV---FKELDQATSKDLAVSTGQLKDVAANNANWFSRAVKVLADIFFV 115

```

Db      61 LGVITGNTGVGVYKALGSGFTKLTDDSDSEIAKGT---KD---SDGFLSKAIDVISGIFT 114
Qy      116 PLIPILVGGGLMANNVLAQDLFGPQSLVEMFPQISGVAMENIMASAPFAFLPVLTG 175
Db      115 PLIGLALGGGKMLGKLMILT---FG-----WLTSSGTYQILYAAADSVFELPLILA 165
Qy      176 FTRATKPGGNEFLGAGIGMAVFPPLTVNGVDVAATMTAGBMPWMSLFGLDVAQAGYQTV 235
Db      166 YTPARKFGANPPPAIAAAGALVYPTMINLFGCAHTPLQIP-----VLMSTYSFSV 217
Qy      236 LPLVSMILATIEKFLHKLRLMGTADELITPVLTLTLTGFLTFEIAIGPAMRWGDLAAG 295
Db      218 IPIILAVWFSLIERLNSKIHAAKTFILPMICMLIPLITFLARGPLGTIFISGLAAG 277
Qy      296 LOGLYDPGGPVGGLFGLVYSPVITGLHQSFPPIEL-ELFNQGSFIFATASMANIAQ 354
Db      278 YTFIYNLSPVIVAGFAPGAFWQVLIIFGIHWGFVPIMINLSRGYRDTMIAMGPSNFAQ 337
Qy      355 AACLAVFLAKSEKLGKLAGAGSVSAVLGITTEPAIFGVNLRMRPFYIGITTAIGGALI 414
Db      338 GASLGVFLTKKPEVNAIAGSALTGFPGITBPSITGVTLKPKPVIASIAIGALGALV 397
Qy      415 ALFDIKAVAGAGFLGVNSIDAPDMV---MPLVCAVTFVIAFGAIAV-----GLYIV 466
Db      398 -----GAAAGSGAANA-IPGILTLPIFIGKGFVGFIL--GIAVAYIISALIGTYF 444
Qy      467 RRRGSDIDPATAPVPDAGTTKAEAPAEPSNDSTIIQAPLGEALIASVSDAMFASK 526
Db      445 ---GYDEADMDGI---APTTEAKETGVE---ABEIVSPIRGINIVPLNEVDEAFSAGL 494
Qy      527 LGSVAIVPTKQOLVSPVSGKIIVAFPSGHAFAVRKAEKDSVNDILMHGPTVNLNGT 586
Db      495 LKGVAIVQEGKLLISPVNGTETAPFTHAIGIRS---DKGVEILLHVGPDTVLQNGK 550
Qy      587 HENPLKQGDENVAGELICEFDIDAIAKAGYEVTTPIVSN 627
Db      551 YFKLVAQGDRLVVGQALIEFDLEAIKADGYDITTPIVTN 591

RESULT 13
E98067
phosphotransferase system enzyme II (EC 2.7.1.69) scra [imported] - Streptococcus pneumo
C.Species: Streptococcus pneumoniae
C.Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C.Accession: E98067
R.Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A.Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;
A.Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A.Reference number: A97872; MUID:21429245; PMID:11544234
A.Accession: E98067
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-627 <KUR>
A.Cross-References: UNIPROT:Q8DMS8, GB:AE007317, PIDN:AAL00370.1, PID:g15459232, GSPDB:C
C.Genetics:
A.Gene: scra
A.Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
C.Keywords: phosphotransferase

Query Match      28.6%; Score 956; DB 2; Length 627;
Best Local Similarity 33.3%; Pred. No. 5.5e-56;
Matches 222; Conservative 141; Mismatches 246; Indels 58; Gaps 14;

```

```

Qy      1 MDHDLAQRILRDIGEDNIVAAAGCATRLRLVLDKTKVDVROSIDDDPDKGTFTGTGM 60
Db      1 MNNGELAKKVIDALGGERENVNSVAHCATRLRWVQDEKINKNEVENLEKVVQCAFNSQ 60
Qy      61 FOIIVPGDVDFVFKF---LDDATSKDIIVSTEQLDKDVANNANMFSAVKVLAIDIFVL 117
Db      61 FOIIVPGDVDFVFKF---LDDATSKDIIVSTEQLDKDVANNANMFSAVKVLAIDIFVL 117

```

```

Db      61 YQIIFGTGTVMNGYDEVVVLGLPTS-----SKDMKAEVAKQGNMFORAIRTEGDFVPI 115
Qy      118 IPIVGGGLMANNVLAQDLFGPQSLVEMFPQISGVAMENIMASAPFAFLPVLTG 177
Db      116 IPIVATGIFMGVIRGIFNALEM-----PLGDFPATYQIILDTAFIILPGLVMS 165
Qy      178 ATKRFKPGNEFLGAGIGMAVFPPLTVNGVDVAATMTAGBMPWMSLFGLDVAQAGYQTV 237
Db      166 TRVREGNPAVAVIYGMMLVSGSLPNAVA---QGGEVTANPFGF-IPVGLQGSVLP 221
Qy      238 VLVSMILATIEKFLHKLRLMGTADELITPVLTLTLTGFLTFEIAIGPAMRWGDLAAG 297
Db      222 AFIIGVAKKFKKAVKVPDVIDLLVMPFVLVMSILGLFVIGFVHVENYILIAK 281
Qy      298 GLYDGGPVGGLFGLVYSPVITGLHQSFPPIEL-ELFNQGSFIF-ATASMANIAQ 356
Db      282 AILSMFPGIGFLIGVQOLIVSGVHIFNLLEVQOLAADHANFNAITTAAMTAQGA 341
Qy      357 CLAVFLAKSEKLGKLAGAGSVSAVLGITTEPAIFGVNLRMRPFYIGITTAIGGALI 416
Db      342 TVAVGVTKNPKLTLAPPAIASFLGITEPAIFGVNLRMRPFYIGITTAIGGALI 401
Qy      417 FDIKAVAGAGFLGVNSIDAPDMV-----MPLVCAVTFVIAFGAIAVGYLYVR 468
Db      402 -----LGLATGNGITIIIPGTMLYGNGQLPQYLLVAVSFGALFTWFGY----- 449
Qy      469 NSIDPDATAPVPDAGTTKAEAE-APAEPSNDSTIIQAPLGEALIASVSDAMFASK 527
Db      450 --EDEVDAVPAAKQAEVAEKEEVAPALQNETLV--TPVVDVVALADVNDVPFSSGAM 505
Qy      528 GSGVAIVPTKQOLVSPVSGKIIVAFPSGHAFAVRKAEKDSVNDILMHGPTVNLNGH 587
Db      506 GGGLVAKVSQGVVYALDAEVSIAFPTHAFGLKTR---NGAEVLHIGIDIVSRNGG 561
Qy      588 FENPLKQGDENVAGELICEFDIDAIAKAGYEVTTPIVSN---YKKTGPNVTYGLGEIRA 644
Db      562 FEAQXAKGNKVAAGVLTGFFDSNKIAAGLDDTTWIVTADVASVAVAT---GSVSK 618
Qy      645 GANLINV 651
Db      619 GDVAVIEV 625

RESULT 14
C95220
trehalose PTS system, IIAB components [imported] - Streptococcus pneumoniae (strain TIC
C.Species: Streptococcus pneumoniae
C.Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C.Accession: C95220
R.Reteljin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A.Authors: Loftus, B.D.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A.Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.
A.Reference number: A95000; MUID:21357209; PMID:11463916
A.Accession: C95220
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-655 <KUR>
A.Cross-References: UNIPROT:Q97NM9; GB:AE005672; PIDN:AK75956.1; PID:g14973388; GSPDB:C
A.Experimental source: strain TIGR4
C.Genetics:
A.Gene: SPI884
A.Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match      28.4%; Score 949; DB 2; Length 655;
Best Local Similarity 34.9%; Pred. No. 1.7e-55;
Matches 234; Conservative 128; Mismatches 253; Indels 56; Gaps 19;

```

```

Qy      7 AQRILRDIGEDNIVAAAGCATRLRLVLDKTKVDVROSIDDDPDKGTFTGTGMFOIIVG 66
Db      8 AKDLDLQALGKENVAVATHCATRMRFLVGDGDKAVKVAIESIPAVKGFITNAGQGVILIG 67

```



```

QY 67 PGDVHVEKELDDATSKDIASVTEOLKDVANNANMFSAVKYIADI FVPLIPILVGGGL 126
DB 68 -NDVPIFYND F--TAVSGIEGVSKAAKSAKSNQNVGVWTLAEFTPIIPALIYGG 125
QY 127 LMAINNTL--VAQDLFGPQSLVEMFPQISGVAEMINMASAPFAPLPVLVGFATKRRGG 184
DB 126 ILGFRNVLEGVHWSMLDKITITESSQFVAGVNHFLMGEALIFQLPVGITWVSRRKMG 185
QY 185 NEPLAGIGMAMVPTLVNGVDVATMTAGEMP--MMSLFGLDVAQAGYQGYTL PVLVS 242
DB 186 SQTIGIYVGLICVSPQLNAYASTPAADIAANNWVNFPGYFTVNRIGYQAQVIPALLAG 245
QY 243 WILATIEKFLHKLMTADFLITPVLTLLTGFLTFIAGPARKWGDLLAHGLOGLYDF 302
DB 246 LSLSYLEIFMRKHIDEVISMIFVFPFLSLIPALILAHVTLGP---IGMTIGGGLSSVLA 301
QY 303 G--GPVG---GLFGLVYSPVITGLHQSFPPIELELF--NQGGSFIFATASMANIAOGAA 356
DB 302 GLTGPKMLFGAIFGALYAPFVITGLHMTNADITQLADAGTALMPMLALSNIAOGSA 361
QY 357 CLAVFPLAK--SEKLKLAGASGSAVLGITEPAIFGVNLRMPFYIGITAAIGALIA 415
DB 362 VFAVYFMRHDERERAQVSLPATISAVLGTEPALFGVNVKTYFPVAGMTGSALAGMLSV 421
QY 416 LFDIKAVALLGAAGFLGVNSIDAPDMVF---LYCAVTVFVIAFGAALAYGLVYRRNGS 471
DB 422 TENVTAASIGIGLPGILISTIQPYMLPFAGTMLVAIVPMLTLP-----FFRKAGL 472
QY 472 IDPDATAAPVAPAGTTKABE---APAEFSNDSTI-----IQAPLTGEALALSSVSAM 521
DB 473 FTK-----TEGDTNLQAEFVAQEBAEFVNSHEVELTSVELISPLTGQVKELSQATDPI 525
QY 522 FASGKLGSVAIVPTKGQVSPVSGKIYVAPSPGHAFVARTKAEDGSNDVILMHIGPTV 581
DB 526 FASGVWGGLVIERPSQGLTSPVNGVTVLFPRTKAIGI--VSDGG--VELLIHIGMDTV 581
QY 582 NLNGTHFNLKKQGDVYKAGELCEPDIDAIKAGVETTPIVSNK---YKTKGVPVNYG 638
DB 582 GLDGKGFESLVQGDHVTVGQQLIRFMDVIKAGLVETTPVITNODAVYATITP-GTTP 640
QY 639 LGEIEAGANLL 649
DB 641 T-TIQAGASLM 650

```

RESULT 15

A99084
 phosphotransferase system enzyme II (BC 2.7.1.69), trep [imported] - Streptococcus pneu
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C:Accession: A99084
 R:Host: J. A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
 e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 Y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: A99084
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1705 <KUR>
 A:Cross-references: UNIPROT:Q8DNI6; GB:AE007317; PIDD:AL00502.1; PID:G15459375; GSPDB:C
 C:Genetics:
 A:Gene: trep
 C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
 C:Keywords: phosphotransferase

Query Match 27.9%; Score 934; DB 2; Length 705;
 Best Local Similarity 34.2%; Pred. No. 1.9e-54;
 Matches 230; Conservative 130; Mismatches 253; Indels 60; Gaps 18;

```

QY 7 AQRILRDIGGEDNIVAAHACATRLRLVLDTKDVDRQSLDDPDLKGFETGFMFOIIVG 66
DB 58 AKDLQAIIGKENVATVTHCATRRFVLGDDKAAKAIASIAVKGTFTINAQFOVIIG 117
QY 67 PGDVHVEKELDDATSKDIASVTEOLKDVANNANMFSAVKYIADI FVPLIPILVGGGL 126
DB 118 -NDVPIFYND F--TAVSGIEGVSKAAKSAKSNQNVGVWTLAEFTPIIPALIYGG 175
QY 127 LMAINNTL--VAQDLFGPQSLVEMFPQISGVAEMINMASAPFAPLPVLVGFATKRRFG 184
DB 126 ILGFRNVLEGVHWSMLDKITITESSQFVAGVNHFLMGEALIFQLPVGITWVSRRKMG 235
QY 176 ILGFRNVLEGVHWSMLDKITITESSQFVAGVNHFLMGEALIFQLPVGITWVSRRKMG 235
DB 185 NEPLAGIGMAMVPTLVNGVDVATMTAGEMP--MMSLFGLDVAQAGYQGYTL PVLVS 242
QY 243 WILATIEKFLHKLMTADFLITPVLTLLTGFLTFIAGPARKWGDLLAHGLOGLYDF 302
DB 236 SQTIGIYVGLICVSPQLNAYASTPAADIAANNWVNFPGYFTVNRIGYQAQVIPALLAG 295
QY 296 LSLSYLEIFMRKHIDEVISMIFVFPFLSLIPALILAHVTLGP---IGMTIGGGLSSVLA 351
DB 303 G--GPVG---GLFGLVYSPVITGLHQSFPPIELELF--NQGGSFIFATASMANIAOGAA 356
QY 352 GLTGPKMLFGAIFGALYAPFVITGLHMTNADITQLADAGTALMPMLALSNIAOGSA 411
DB 357 CLAVFPLAK--SEKLKLAGASGSAVLGITEPAIFGVNLRMPFYIGITAAIGALIA 415
QY 412 VFAVYFMRHDERERAQVSLPATISAVLGTEPALFGVNVKTYFPVAGMTGSALAGMLSV 471
DB 416 LFDIKAVALLGAAGFLGVNSIDAPDMVF---LYCAVTVFVIAFGAALAYGLVYRRNGS 471
QY 472 IFNVTAASIGIGLPGILISTIQPYMLPFAGTMLVAIVPMLTLP-----FFRKAGL 522
DB 472 IDPDATAAPVAPAGTTKABE---APAEFSNDSTI-----IQAPLTGEALALSSVSAM 521
QY 523 FTK-----TEGDTNLQAEFVAQEBAEFVNSHEVELTSVELISPLTGQVKELSQATDPI 575
DB 522 FASGKLGSVAIVPTKGQVSPVSGKIYVAPSPGHAFVARTKAEDGSNDVILMHIGPTV 581
QY 576 FASGVWGGLVIERPSQGLTSPVNGVTVLFPRTKAIGI--VSDGG--VELLIHIGMDTV 631
DB 582 NLNGTHFNLKKQGDVYKAGELCEPDIDAIKAGVETTPIVSNK---YKTKGVPVNYG 636
QY 632 GLDGKGFESLVQGDHVTVGQQLIRFMDVIKAGLVETTPVITNODAVYATITGTP 691
DB 637 YGLIEAGANLL 649
QY 692 ----TIQAGASLM 700

```

Search completed: March 7, 2005, 22:16:51
 Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:57:13 ; Search time 180 Seconds
(without alignments)
1880.469 Million cell updates/sec

Title: US-10-019-284B-2

Perfect score: 3342

Sequence: 1 MDHKLQRIQLRDIGGEDNI.....TEAGANLVAKKEAVPATP 661

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3310	99.0	661	2	Q8NMD6
2	1609.5	48.2	651	1	PTSA_PDPPE
3	1607.5	48.1	651	1	Q88ZV9
4	1508	45.1	650	2	Q9S6S6
5	1478.5	44.2	664	1	PTSA_STPMU
6	1457	43.6	649	2	Q7WMP7
7	1446.5	43.3	647	2	Q74H19
8	1345.5	40.3	643	2	Q82YR5
9	1264	37.8	534	2	Q93ML1
10	1227	36.7	480	1	PTSB_STPMU
11	1186	35.5	481	2	Q8CN82
12	1183	35.4	480	2	Q8NVT5
13	1183	35.4	480	2	Q6G6U2
14	1179	35.3	480	2	Q6G6U2
15	1178	35.2	480	2	Q99RQ0
16	1178	35.2	480	2	Q7A3V6
17	1160.5	34.7	492	2	Q6SUG9
18	1100	32.9	474	1	PTSB_PSMU
19	1098.5	32.9	472	2	Q9L8G6
20	1093	32.7	479	2	Q9KLT8
21	1089	32.0	479	1	PTSB_VIBAL
22	1013.5	30.3	470	2	Q6SMI3
23	1006	30.1	470	1	PTTB_BACSU
24	999	29.9	470	1	Q8G1N4
25	979.5	29.3	480	2	Q831B4
26	974	29.1	475	2	Q81V69
27	964.5	28.9	475	2	Q81V69
28	963.5	28.8	475	2	Q6HNT2
29	961.5	28.8	475	2	Q73DL3
30	960	28.7	627	2	Q97PB8
31	957	28.6	617	2	Q8Y904

32	956.5	28.6	475	2	Q63G14	Q63G14 bacillus ce
33	936.5	28.6	480	2	Q9F8X3	Q9F8X3 pseudomonas
34	956	28.6	627	2	Q8DNS8	Q8DNS8 streptococ
35	954.5	28.6	475	2	Q81H25	Q81H25 bacillus ce
36	951.5	28.5	627	2	Q8N2J7	Q8N2J7 streptococ
37	949	28.4	655	2	Q97NW9	Q97NW9 streptococ
38	946	28.3	614	2	Q65D36	Q65D36 bacillus 11
39	937.5	28.1	627	2	Q8K524	Q8K524 streptococ
40	935.5	28.0	620	2	Q99Y91	Q99Y91 streptococ
41	934	27.9	619	2	Q65D51	Q65D51 bacillus 11
42	934	27.9	705	2	Q8DNI6	Q8DNI6 streptococ
43	933.5	27.9	475	2	Q99WC9	Q99WC9 staphylococ
44	933.5	27.9	475	2	Q7A7D1	Q7A7D1 staphylococ
45	933	27.9	674	2	Q8K5K9	Q8K5K9 streptococ

ALIGNMENTS

RESULT 1

Q8NMD6 PRELIMINARY; PRT; 661 AA.

AC Q8NMD6; Q6M2J8; 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 28, Last annotation update)

DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)

DE Phosphotransferase system ITC components, Glucose/malose/N-acetylglucosamine-specific (EC 2.7.1.69) (ENZYME II SUCROSE PROTEIN).

GN Name=ptss: OrderedLocustNames=Cg12642, CG2925;

OS Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

OX NCBI_TaxID=1718;

[1]

RP SEQUENCE FROM N.A.

RA Nakagawa S.;

RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";

RT Submitted (May-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;

RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M., Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Galglat L., Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B., McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A., Rey D.A., Ruckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I., Tauch A.;

RA "The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins.";

RT J. Biotechnol. 104:5-25(2003).

DR EMBL: AP005282; BAC00361; --

DR EMBL: BX927155; CAF21304.1; --

DR HSSP: P45618; 2GPR.

DR GO: GO:0016020; C:membrane; IEA.

DR GO: GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. . ; IEA.

DR GO: GO:0005351; F:sugar porter activity; IEA.

DR GO: GO:0016740; F:transferase activity; IEA.

DR GO: GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . ; IEA.

DR GO: GO:0006810; P:transport; IEA.

DR GO: GO:0011055; DUP_hybrid motif.

DR InterPro: IPR011955; Dup_hybrid motif.

DR InterPro: IPR001996; Ptrans_EIIB.

DR InterPro: IPR003352; Ptrans_EIIB.

DR InterPro: IPR001127; Pts_EIIA.

DR InterPro: IPR010973; Pts-II_BC_sucr.

DR Pfam: PF00358; Pts_EIIA_1; 1.

DR Pfam: PF00367; Pts_EIIB; 1.

DR Pfam: PF02378; Pts_EIIC; 1.

DR ProDom: PD001476; Ptrans_EIIB; 1.

DR ProDom: PD002243; Pts_EIIA; 1.

DR TIGRFAMS; TIGR00830; PTB; 1.
 DR TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.
 DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
 KM Complete proteome; Transferrase.
 SQ SEQUENCE 661 AA; 69148 MW; COSESLC7833B9F74 CRC64;

Query Match 99.0%; Score 3310; DB 2; Length 661;
 Best Local Similarity 98.9%; Pred. No. 3.3e-191;
 Matches 654; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDHKLARILRDIGGEINIVAAACATRLRLVLTQTDVDRQSDDDPDLKGFETGGM 60
 DB 1 MDHKLARILRDIGGEINIVAAACATRLRLVLTQTDVDRQSDDDPDLKGFETGGM 60
 QY 61 FOIIVGPDVHVFKELDATSKDIASVTEQDKDVANNANMFSAVKYADIIVPLPI 120
 DB 61 FOIIVGPDVHVFKELDATSKDIASVTEQDKDVANNANMFSAVKYADIIVPLPI 120
 QY 121 LVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINLMASAPFAFLPVLVGFTATK 180
 DB 121 LVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINLMASAPFAFLPVLVGFTATK 180
 QY 181 REGGNEFLGAGIGMAVPEPTLVNGVDVATMTAGEMPMSLFGLDVAQGYQGYVLPYV 240
 DB 181 REGGNEFLGAGIGMAVPEPTLVNGVDVATMTAGEMPMSLFGLDVAQGYQGYVLPYV 240
 QY 241 VSMILATIEKFLHKRLMTADELITPVLTLLTGLFTFIAIGPAMRWGDLIAHGLQGLY 300
 DB 241 VSMILATIEKFLHKRLMTADELITPVLTLLTGLFTFIAIGPAMRWGDLIAHGLQGLY 300
 QY 301 DRGGPVGILLFGLVSPYITITGLHOSFPRIELELFQGGSFIFATASMANIAQGAACLV 360
 DB 301 DRGGPVGILLFGLVSPYITITGLHOSFPRIELELFQGGSFIFATASMANIAQGAACLV 360
 QY 361 FFLAKSEKIKGLAGASGAVIGITEPAIFGVNLRIRMFYIGIGTAIGALIALFDIK 420
 DB 361 FFLAKSEKIKGLAGASGAVIGITEPAIFGVNLRIRMFYIGIGTAIGALIALFDIK 420
 QY 421 AVALGAAGFLGVSDIDPMVMFLVCAVTFYIAFGAIAIYGLVYRNGSIDPDATAAP 480
 DB 421 AVALGAAGFLGVSDIDPMVMFLVCAVTFYIAFGAIAIYGLVYRNGSIDPDATAAP 480
 QY 481 VPAGTTKARAEAPAESNDSTIIQALTEBAIALSSVSAMPASGLSGSVAIVPTKGL 540
 DB 481 VPAGTTKARAEAPAESNDSTIIQALTEBAIALSSVSAMPASGLSGSVAIVPTKGL 540
 QY 541 VSPVSGKIIVAFPSGHAFAVRTKAEDEGSNVLDIMHIGFDTVNLNGTHFNPPLKQGEVKA 600
 DB 541 VSPVSGKIIVAFPSGHAFAVRTKAEDEGSNVLDIMHIGFDTVNLNGTHFNPPLKQGEVKA 600
 QY 601 GELLCEFDIDAIQAAGYEVTTPIVGSNNYKKTGPVNTYGGIETIAGANILNAKKEAVPAT 660
 DB 601 GELLCEFDIDAIQAAGYEVTTPIVGSNNYKKTGPVNTYGGIETIAGANILNAKKEAVPAT 660
 QY 661 P 661
 DB 661 P 661

RESULT 2

PTSA_PDPPE STANDARD; PRT; 651 AA.
 AC P43470;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE PTS system, sucrose-specific IIBAC component (EIIABC-Scr) (Sucrose-
 de permease IIBAC component) (Phosphotransferase enzyme II, ABC
 component) (EC 2.7.1.69) (EII-Scr).
 GN Name=scrA;
 OS Pedicoccus pentosaceus.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Pedicoccus.
 NCBI_TaxID=1255;

EN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEP1.0;
 RA Leenhouts K.K.J., Bolhuis A.A., Kok J.J., Venema G.G.;
 RT "The sucrose and raffinose operons of *Pedococcus pentosaceus*
 RP1.0";
 RL Submitted (Apr-1994) to the EMBL/GenBank/DBJ databases.
 CC -1 FUNCTION: This is a component of the phosphoenolpyruvate-dependent
 CC sugar phosphotransferase system (PTS), a major carbohydrate active
 CC -transport system. The IID domain contains the sugar binding site
 CC and the transmembrane channel; the IIA domain contains the primary
 CC phosphorylation site (the donor is phospho-HPr); IIB transfers its
 CC phosphoryl group to the IIB domain which finally transfers it to
 CC the sugar.
 CC -1 CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 CC histidine + sugar phosphate.
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1 SIMILARITY: Contains 1 PTS EIIA domain.
 CC -1 SIMILARITY: Contains 1 PTS EIIA domain.
 CC -1 SIMILARITY: Contains 1 PTS EIIA domain.
 CC -1 SIMILARITY: Contains 1 PTS EIIA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC EMBL; Z32771; CAA83668.1; -;
 CC EMBL; L32093; AAB25567.1; -;
 CC PIR; S44257; S44257.
 CC HSSP; P20166; IGBR.
 DR InterPro; IPR011055; Dup_hybrid_motif.
 DR InterPro; IPR001996; Pterans_EIIB.
 DR InterPro; IPR003352; Pterans_EIIC.
 DR InterPro; IPR001127; PTS_EIIA.
 DR InterPro; IPR010973; PTS_II_BC_sucr.
 DR Pfam; PF00359; PTS_EIIA_1; 1.
 DR Pfam; PF00367; PTS_EIIB_1.
 DR Pfam; PF02378; PTS_EIIC_1.
 DR ProDom; PD001476; Pterans_EIIB; 1.
 DR ProDom; PD002243; PTS_EIIA_1.
 DR TIGRFAMS; TIGR00826; EIIA_glc; 1.
 DR TIGRFAMS; TIGR00830; PTB; 1.
 DR TIGRFAMS; TIGR01995; PTS-II-ABC-beta; 1.
 DR TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.
 DR TIGRFAMS; TIGR01992; PTS-II-BC-Tre; 1.
 DR PROSITE; PS00371; PTS_EIIA_1; 1.
 DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
 KM Phosphorylation; Phosphotransferase system; Sugar transport;
 KW Transferrase; Transmembrane.
 FT DOMAIN 1 40 EIIA.
 FT DOMAIN 111 418 EIIA.
 FT DOMAIN 510 651 EIIA.
 FT MOD RES 25 25 Phosphocysteine (By similarity).
 FT MOD RES 324 324 Phosphohistidine (By similarity).
 FT MOD RES 562 562 Phosphohistidine (By similarity).
 SQ SEQUENCE 651 AA; 68454 MW; C87BA09D550A77F8 CRC64;

Query Match 48.2%; Score 1609.5; DB 1; Length 651;
 Best Local Similarity 49.6%; Pred. No. 8.8e-89;
 Matches 331; Conservative 114; Mismatches 196; Indels 27; Gaps 8;

QY 1 MDHKLARILRDIGGEINIVAAACATRLRLVLTQTDVDRQSDDDPDLKGFETGGM 60
 DB 1 MDHKLARILRDIGGEINIVAAACATRLRLVLTQTDVDRQSDDDPDLKGFETGGM 60
 QY 61 FOIIVGPDVHVFKELDATSKDIASVTEQDKDVANNANMFSAVKYADIIVPLPI 118
 DB 60 YOIIVGPDVHVFKELDATSKDIASVTEQDKDVANNANMFSAVKYADIIVPLPI 118
 QY 119 PIVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINLMASAPFAFLPVLVGFTA 178

```

Db      119 PALVAGGLMALNNVLTAEHLFMKXSVSEVYFGLKGIAMINAMASAPFTPIILLGFGSA 178
Qy      179 TRRFGNEFLGAGIGAMVFTLVNGYDVAAITMTGEMMMSLFGLDVAQAGYGVLPV 238
Db      179 TRRFGNEFLGAGIGAMVFTLVNGYDVAAITMTGEMMMSLFGLDVAQAGYGVLPV 238
Qy      239 LVVSMILATIEKFLHRLMGTADFLITPVLTLLTGFLTFIAGPMRWGDLAHLQ 298
Db      239 LGVATLTLAEKFEFHKIHGADFPTTTPFAIVITGFLFTTVGPVLKRVSDALNGLVG 298
Qy      299 LVDFGPGVGGLLFGLVSPVITVTLGHQSPPIELFLF---NQGSGFIFATASMANIAG 354
Db      299 LVNSGIMGMGIFGLLYSAIVITGLHQTFPALETQLANVAKTGGSFIFPVSAMNIGG 358
Qy      355 AACLAVFLAKSEKLGLAGASGVAVGITTEPAIFGVNLRMPYITIGTAAIGALIT 414
Db      359 AATLTAIFATKQOKOKALTSAGVSALGITTEPAIFGVNLRMPYITIGTAAIGALIT 418
Qy      415 ALFDIKAVNLGAAGFLGVVSDAPDMVMEFLVCVVTFVIAFGAATAVGLVVRNGSIDP 474
Db      419 GLFHLVSAMGPASVIGFISISKSPAPMLSAVISFVAFTPTFY---AKRTIGDDR 474
Qy      475 DATAAPVPAGITKAEAPAEPSNDSTIIQAPLTGEAIALSSVSDAMPASGKLGSGVALV 534
Db      475 DVKSPAPFTSTV-----INVNDEIISAPVTGASESLKQVNDQVFGABIMKGAALV 525
Qy      535 PRKQGVSPVSGKLVVAPSPGSAFAVRTAEEDGSNDVILMHITGPTVNLNGHFNPLKQ 594
Db      536 PSSDQVVAADAVITVYTDHSHAYGKTTA---GAEIILIHGLPTVNLNGHFNPTNNQ 581
Qy      595 GEVAKAGELLCEFDIDAKAGYEVTTPVIVSNYKKGTPVNTYVGEIEBAGNL--LVNA 652
Db      592 GDTVAGDGLGTFDDIALKAANYDPTVMTLVNTANTYANVERLKTNNVAGSGLVALTAP 641
Qy      653 KKEAVPAT 660
Db      642 AASSVAAT 649

RESULT 3
Q88ZV9 PRELIMINARY; PRT; 651 AA.
AC Q88ZV9;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DN Name=pts1BCA; OrderedlocusNames=lp_0185;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxId=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFSL;
RX MEDLINE=22480296; PubMed=1256556; DOI=10.1073/pnas.0337704100;
RA Kleebeezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kluysers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fliers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Sleszen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFSL.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11990-11995(2003).
DR EMBL; AL935252; CAD62855.1; -.
DR HSB; P20166; 1GPR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. . .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup_hybrid_motif.

```

```

DR InterPro; IPR001996; Putane EIIB.
DR InterPro; IPR003352; Putane EIIC.
DR InterPro; IPR001327; Putane EIIC.
DR InterPro; IPR010973; PTS_II_EIIA.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD001476; Putane EIIB; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMs; TIGR00830; PTBA; 1.
DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_Cys; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 651 AA; 68514 MW; CBFBS35EDDB73DD6 CRC64;

Query Match 48.1%; Score 1607.5; DB 2; Length 651;
Best Local Similarity 49.3%; Pred. No. 1.2e-88;
Matches 332; Conservative 114; Mismatches 189; Indels 39; Gaps 9;

Qy      1 MDHKLADRLRDIGDENIVAAACATRLRLVLTQTDVDRQSLDDDDPDLKGTETGGM 60
Db      1 MNHGEVADRVLNATL-GKNNIQAAACATRLRLVTKDESKIDQALDDADVKGTEETNGQ 59
Qy      61 FOIIVPGDVDFHVEK-----LDDATSKDIAVSTEQDKDVA--NNAMWFSBAVKVLAD 112
Db      60 YQIITGGPDVDKVTYALAKTGLKEATPDDI-----KAVAAAGQKKPLMDPLKVLSD 112
Qy      113 IFVPLIPLVGGGLLMAINNVLAQDLFQPSLVEMFPQISGVAMINLMASAPAPLPV 172
Db      113 IFIPVLPVAVGGGLMALNNVLTAEHLFMKXSVSEVYFGLKGIAMINAMASAPFTPIPI 172
Qy      173 LVGFATATGFGNEFLGAGIGAMVFTLVNGYDVAAITMTGEMMMSLFGLDVAQAGY 232
Db      173 LGFSATATGFGNPNLGAITMGIMVLPVLNGYSVAITTAAGKMYVNVFGHLVAQAGY 232
Qy      233 GTVLPVLVSMILATIEKFLHRLMGTADFLITPVLTLLTGFLTFIAGPMRWGDL 292
Db      233 GQVLPVLVAVIYTLAEKFEFHKIHGADFPTTTPFAIVITGFLFTTVGPVLKRVSDAL 292
Qy      299 LVDFGPGVGGLLFGLVSPVITVTLGHQSPPIELFLF---NQGSGFIFATASMANIAG 354
Db      293 TNGLVGLVNTGIMGMGIFGLLYSAIVITGLHQTFPALETQLANVAKTGGSFIFPVSAM 352
Qy      349 ANIAGAACLAVFLAKSEKLGLAGASGVAVGITTEPAIFGVNLRMPYITIGTAAIGALIT 408
Db      353 ANIAGAAATLTAIFATKQOKOKALTSAGVSALGITTEPAIFGVNLRMPYITIGTAAIGALIT 412
Qy      409 ICGALIALFDIKAVNLGAAGFLGVVSDAPDMVMEFLVCVVTFVIAFGAATAVGLVVR 468
Db      413 IASAFGLFHLVSAMGPASVIGFISISKSPAPMLSAVISFVAFTPTFY---AKR 468
Qy      469 NGSIDPDATAAPVPAGITKAEAPAEPSNDSTIIQAPLTGEAIALSSVSDAMPASGKLG 528
Db      469 TLGDRODVKSPAPFTSTV-----INVNDEIISAPVTGASESLKQVNDQVFGABIMG 519
Qy      529 SGVALVPTKQGVSPVSGKLVVAPSPGSAFAVRTAEEDGSNDVILMHITGPTVNLNGHFN 588
Db      520 KGAALVPADQVVAADAVITVYTDHSHAYGKTTA---GAEIILIHGLPTVNLNGHFN 575
Qy      589 NPLKQGEVAVAGELLCEFDIDAKAGYEVTTPVIVSNYKKGTPVNTYVGEIEBAGNL 648
Db      576 TTNVQKGTVAGDGLGTFDDIALKAANYDPTVMTLVNTANTYANVERLKTNNVAGSGL 635
Qy      649 LNVAKKA--VPAT 660
Db      636 VALTEPTASSVAAT 649

RESULT 4
Q9S6S6 PRELIMINARY; PRT; 650 AA.
AC Q9S6S6;

```

DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Enzyme II sucrose protein (EC 2.7.1.69).
 GN Name=sacB;
 OS Lactococcus lactis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1358;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99173919; PubMed=10074089;
 RA Luesink E.J., Marugg J.D., Kuipers O.P., De Vos W.M.;
 RT "Characterization of the divergent sacB and sacA operons, involved
 in sucrose utilization by Lactococcus lactis.";
 RL J. Bacteriol. 181:1924-1926(1999).
 DR EMBL; 297015; CAB09690.1; -.
 DR HSSP; P45618; 2GPR.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0006982; P:protein-N(PI)-phosphohistidine-sugar phospho. . .; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR011055; Dup hybrid motif.
 DR InterPro; IPR001996; Ptrans_EIIB.
 DR InterPro; IPR003352; Ptrans_EIIB.
 DR InterPro; IPR001127; Pts_EIIA.
 DR InterPro; IPR010973; Pts_II_BC_sucr.
 DR Pfam; PF00358; Pts_EIIA.1.
 DR Pfam; PF00367; Pts_EIIB.1.
 DR Pfam; PF02378; Pts_EIIB.1.
 DR ProDom; PD001476; Ptrans_EIIB.1.
 DR ProDom; PD002243; Pts_EIIA.1.
 DR TIGRFAMs; TIGR00830; PTA; 1.
 DR TIGRFAMs; TIGR01996; Pts-II-BC-sucr; 1.
 DR PROSITE; PS00371; Pts_EIIA.1; 1.
 DR PROSITE; PS01035; Pts_EIIB_Cys.1.
 KM Transferase.
 SQ SEQUENCE 650 AA; 69636 MW; 250E26F7664D204 CRC64;
 Query Match 45.1%; Score 1508; DB 2; Length 650;
 Best Local Similarity 46.9%; Pcred. No. 1.1e-82;
 Matches 313; Conservative 123; Mismatches 192; Indels 40; Gaps 10;
 QY 1 MDHKLQRIKIDIGEDNIVAAHCATRLRLVLTQYKQDVQSDDDPDLKGTETGM 60
 DB 1 NMHKQVARIINAV-GRDNIQGRHCATRLRLVLTQYKQDVQSDDDPDLKGTETGAQ 59
 QY 61 PGIITGGGVVDHVPKRLDQATSKDIASVTEOKDVAN--NANWFSBAKVLADI FVPLI 118
 DB 60 YQIIVGPGVNTVYEFKLTGISER-STADLKEIAGSKKQNPVMAVXLSDIFVPLI 118
 QY 119 PIVGGGLLMAINNVLVADLDFPQSLVEMPQISGVAMINIMASAPFAPLVVGFTA 178
 DB 119 PALVAGGLLMAINNVLVADLDFPQSLVEMPQISGVAMINIMASAPFAPLVVGFTA 178
 QY 179 TTRFGGNEFLGAGIGAMVFPVLVNGYDVAATMTAGKPMWSLFGLDVAQAQGYGTLPV 238
 DB 179 TTRFGGNEFLGAGIGAMVFPVLVNGYDVAATMTAGKPMWSLFGLDVAQAQGYGTLPV 238
 QY 229 LTVSMVLTITKEFLKRLMGTDPLITPVLTLTLGFLFIAGPMRMVGGDLAHGQ 298
 DB 229 LTVSMVLTITKEFLKRLMGTDPLITPVLTLTLGFLFIAGPMRMVGGDLAHGQ 298
 QY 239 IGVAAILAKLERFFPHKYLNDADFTFTPLSLVITIGFLFTTVGPKLRVNSGLDGLVG 298
 DB 239 IGVAAILAKLERFFPHKYLNDADFTFTPLSLVITIGFLFTTVGPKLRVNSGLDGLVG 298
 QY 239 LVDFFGPGVGLFGVLVSPVITVGLHQSPPEPTELELFNO-----GGSFIFATASMANIA 352
 DB 239 LVDFFGPGVGLFGVLVSPVITVGLHQSPPEPTELELFNO-----GGSFIFATASMANIA 352
 QY 299 LVNTIGALGMVFGGYSAIVTGLHQSPPEPTELELFNO-----GGSFIFATASMANIA 358
 DB 299 LVNTIGALGMVFGGYSAIVTGLHQSPPEPTELELFNO-----GGSFIFATASMANIA 358
 QY 333 QGACCLAVFFLAKSEKLGKLAGAGVSAVLGITTEPAIFGVNRLRMPPFTIGTALGCA 412
 DB 333 QGACCLAVFFLAKSEKLGKLAGAGVSAVLGITTEPAIFGVNRLRMPPFTIGTALGCA 412
 QY 359 QGATFFAIFVTKIKTKTALAAPAGVASALIGITEPALFQINIKTKYFPFIALGASAGSL 418
 DB 359 QGATFFAIFVTKIKTKTALAAPAGVASALIGITEPALFQINIKTKYFPFIALGASAGSL 418
 QY 413 LIALPDIKVALGAAGFLGVSDIDAPDMVFLVCAVVFVIAFGAIALVGLVVRNGSI 472
 DB 413 LIALPDIKVALGAAGFLGVSDIDAPDMVFLVCAVVFVIAFGAIALVGLVVRNGSI 472

DB 419 FMGLFHLVALSGLGILGIFISIKAGYVLQGMISIFSLAFLVFTSYGRMAK--SI 476
 QY 473 DDDATAAPVPAATTKAEAEAPAEFNSDSTIIQ-----APLTGEALISSVSDAMFAS 524
 DB 477 TRE-----KNQNA-TTQYQPEKVIIDPVKSGELAPINGFVPLSDVSDPVRSK 525
 QY 525 GKLSGVAIVPTKQOLVSPVSGKIVVAPPSGHAARVTRKADSGNVDIMHIGFDVNLN 584
 DB 526 EIMGGGIAIKPSGSELPADGEIITIAETHAAGIKTR---NGGEVLHIGIDVSMN 581
 QY 585 GTHFNPILKQGDDEVYAGELCEFPIDAIKAGYEVTPFIVSN--YKKTGPVNTYGLGE 641
 DB 582 GNGFQNVKQVQKAKADLLSPDKERIKSGLDPTVIIVITNAGSYNEIIPLEEN--VD 639
 QY 642 IEAGANLL 649
 DB 640 IKVGEKIL 647
 RESULT 5
 ID PTSA STRMU STANDARD; PRT; 664 AA.
 AC P12655;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE PTS system, sucrose-specific IIAB component (EIIB-Scr) (sucrose-
 permease IIAB component) (Phosphotransferase enzyme II, ABC
 component) (EC 2.7.1.69) (EII-Scr).
 GN Name=scrA; OrderedLocusNames=SMU.1841;
 OS Streptococcus mutans
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=89123027; PubMed=2536656;
 RA Sato Y., Foy F., Jacobson G.R., Kuramitsu H.K.;
 RT "Characterization and sequence analysis of the scrA gene encoding
 enzyme IICer of the Streptococcus mutans phosphoenolpyruvate-dependent
 sucrose phosphotransferase system.";
 RL J. Bacteriol. 171:263-271(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 RN [3]
 RP SEQUENCE OF 639-664 FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=93329360; PubMed=8336109;
 RA Sato Y., Yamamoto Y., Kizaki H., Kuramitsu H.K.;
 RT "Isolation, characterization and sequence analysis of the scrA gene
 encoding fructokinase of Streptococcus mutans.";
 RL J. Gen. Microbiol. 139:921-927(1993).
 CC -I- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
 sugar phosphotransferase system (PTS), a major carbohydrate active
 transport system. The IICD domains contain the sugar binding site
 and the transmembrane channel; the IIA domain contains the primary
 phosphorylation site (the donor is phospho-HPr); IIA transfers its
 phosphoryl group to the IIB domain which finally transfers it to
 the sugar.
 CC -I- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 histidine + sugar phosphate.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

```

CC -1- SIMILARITY: Contains 1 PTS_EI1A domain.
CC -1- SIMILARITY: Contains 1 PTS_EI1B domain.
CC -1- SIMILARITY: Contains 1 PTS_EI1C domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL: M22711; AAA26971.1; -.
CC EMBL: AE015011; AAN59464.1; -.
CC EMBL: D13175; BAA02466.1; -.
CC PIR: B32243; B32243.
CC HSSP: P08837; 1GGR.
CC InterPro: IPR011055; Dup hybrid motif.
CC InterPro: IPR001996; Ptrans_EI1B.
CC InterPro: IPR003352; Ptrans_EI1C.
CC InterPro: IPR001127; PTS_EI1A.
CC InterPro: IPR010973; PTS_EI1A_1; 1.
CC Pfam: PF00358; PTS_EI1A_1; 1.
CC Pfam: PF02378; PTS_EI1B; 1.
CC ProDom: PD001476; Ptrans_EI1B; 1.
CC ProDom: PD002243; PTS_EI1A; 1.
CC TIGRfam: TIGR00826; E1B_glc; 1.
CC TIGRfam: TIGR00830; PTBA; 1.
CC TIGRfam: TIGR01995; PTS-II-ABC-beta; 1.
CC TIGRfam: TIGR01996; PTS-II-BC-sucr; 1.
CC TIGRfam: TIGR01992; PTS-II-BC-tre; 1.
CC PROSITE: PS00371; PTS_EI1A_1; 1.
CC PROSITE: PS01035; PTS_EI1B_Cys; 1.
CC Complete proteome; Phosphorylation; Phosphotransferase system;
CC Sugar transport; Transferase; Transmembrane.
CC KW DOMAIN 1 40 EI1B.
CC FT DOMAIN 111 420 EI1C.
CC FT MOD_RES 53 664 EI1A.
CC FT MOD_RES 25 25 Phosphocysteine (By similarity).
CC FT MOD_RES 331 331 Phosphohistidine (By similarity).
CC FT MOD_RES 585 585 Phosphohistidine (By similarity).
CC SQ SEQUENCE 664 AA; 69988 MW; 8095633281A9A1 CRC64;

Query Match 44.2%; Score 1478.5; DB 1; Length 664;
Best Local Similarity 46.0%; Pred. No. 7e-81;
Matches 311; Conservative 127; Mismatches 203; Indels 35; Gaps 13;

QY 1 MHHKDLAQRILRDIGEDENIVAAAHGATRLRLVLDKTDVDRQSLDDDDPKGTPTGGM 60
DB 1 MYSKVASSEVITAV-GKDNLVAAAHGATRLRLVLDKDSKVDQALDKADVKGTPTGGM 59
QY 61 FOIIVGGDVHDVFKELDDATSKDIIVSTEOLKDVANNA--NMFSAVKVLADIFVPLI 118
DB 60 YQVIIGPGDVNFYDEIRIKQTGL-TEVSTDDLKKIAGSKKFNPIWALKLSLDIVPII 118
QY 119 PLVVGGLMANNVAVADLFGPQSLVEMFPOISVAENIMASAPAPFVLVGFPA 178
DB 119 PALVVGGLMALNNFLTSGLFETKSLVVOOFLIKSSDMDQLMSAAPWFLPILVIGISA 178
QY 179 TRRFGNEFLAGIGIAGMAVFTLVNGVDVAA-----TWTAGMP-MWGLFGLDVQAQY 231
DB 179 AGRFGANQFLGASIGIMVAPGANIIGLAANAIPISKATITAGYGFPMIIFGLHTQASY 238
QY 232 OCTVLVPLVSVILATIEFKRLKRLMGTADEFITPVLTLLTGFLTIAIGPAMRWGL 291
DB 232 TVQVPLVLAVWLLSTIEKFPKRLPSAVDFETPLSLVINGFLTFVIYIGPMKEVSDM 298
QY 292 LAHGQGLVDFGPGVGGLLFGLVYSPIVTGLHQSPPLELE---FNGS---GSTFAT 345
DB 292 LINGVIMVLDTTGFGMGVFGALYSPVMTGLHQSPPLELE---FNGS---GSTFAT 358
QY 346 ASMAIAGCAACLAFFLAKSEKLGKLAGASGVAVLGITPEAIFGVNRLRMPFIYIG 405

```

```

DB 359 ASMAIAGCAAFPAIYFLTKDKKKMGLSSSGVSALGITPEALGVNLRKPFPCAL 418
QY 406 TTAIGGALLALPDIKAVALGAAGFLGVSIDAPDMWFLVCAVTFVIAFGAIAVGLY 465
DB 419 GSASMAAIIAGLQVAVVAGSAGFLGSLIKASIPFYVCELISPAIAFATYVG--- 475
QY 466 VRRNGSIDPDAAPAPPACTTAEAAPAEFNSDST-----IIQPLTGEAIALSVS 518
DB 476 --KTVAVDFAAEAVAEAEIIEVQ-EIPEEAASANKAQVTDEVLAAPLAGEAVELTSV 532
QY 519 DAMFASGKSGVAIVPTKGLVSPVSGKIIVAPSPGHAFAVTRAEAGSDNVDILMHGP 578
DB 533 DVFSEANQKGIATKPSNTYVAPVDGTVOAIFDGHAYGI--KSDNGA--ELIHIGI 588
QY 579 DTVNINGTHFNPLKQGEDEKAGELLCEFDIDAIKAAGYVTTPIVNSYKKTGPVNTY 637
DB 589 DTVSMGKGFEGQVQADQIKKGDVLTGTFDSQKIAEAGLDNTTMTFVTADYASVETLA 648
QY 638 GLGEIEAGANLNVAK 653
DB 649 SSGTVAVGDLSLEVRK 664

RESULT 6
ID Q7WMP7 PRELIMINARY; PRT; 649 AA.
AC Q7WMP7;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Sucrose PTS transporter (EC 2.7.1.69).
GN Name=sctra;
OS Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OC NCBI_Taxid=1579;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCFW;
RX PubMed=12847288; DOI=10.1073/pnas.1332765100;
RA Barrangou R., Altermann E., Hutkins R., Cano R., Klaenhammer T.R.;
RT "Functional and comparative genomic analyses of an operon involved in
RT fructooligosaccharide utilization by Lactobacillus acidophilus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8957-8962(2003).
DR EMBL: AY177419; AA038866.1; -.
DR HSSP: P20166; 1AX3.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008982; P:protein-N(PI)-phosphohistidine-sugar phospho. .; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR011055; Dup hybrid motif.
DR InterPro: IPR001996; Ptrans_EI1B.
DR InterPro: IPR001127; PTS_EI1A.
DR InterPro: IPR010973; PTS_EI1A_1; 1.
DR Pfam: PF00358; PTS_EI1A_1; 1.
DR Pfam: PF00367; PTS_EI1B; 1.
DR ProDom: PD001476; Ptrans_EI1B; 1.
DR ProDom: PD002243; PTS_EI1A; 1.
DR TIGRfam: TIGR00830; PTBA; 1.
DR TIGRfam: TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE: PS00371; PTS_EI1A_1; 1.
DR PROSITE: PS01035; PTS_EI1B_Cys; 1.
KW Transferase.
SQ SEQUENCE 649 AA; 69446 MW; B437CA24D00507B CRC64;

Query Match 43.6%; Score 1457; DB 2; Length 649;
Best Local Similarity 45.0%; Pred. No. 1.3e-79;
Matches 290; Conservative 122; Mismatches 191; Indels 42; Gaps 10;

QY 1 MHHKDLAQRILRDIGEDENIVAAAHGATRLRLVLDKTDVDRQSLDDDDPKGTPTGGM 60

```

```

Db      1 MDHKVAERVKDV--GRDNIITAGACATRLRLVLTDDSKVQKALDNDPDKVGTGFKTNGQ 59
Qy      61 FOIIVGPDVHVFEKELDDATS--KQIAVSTBQLKDVNA--NNMFSRAVKVLADIIVPL 117
Db      60 YQVILIGPDVNVVDEFTIKITGLKEL--STDLDKVAABEGQKQKIMPIFKILSIIFPI 117
Qy      118 IPIIVGGGLMANNVLAODLFGPQSLVEMPPQISGVAEMINIMASAPFALPVLVGFT 177
Db      118 IPALVAVGLMALNNFLTSPGLFGKSVYQMAPNVAIGISSMIVQMSAPFIEMPIVMS 177
Qy      178 ATKRGSGHEFLGAGIGMAMVPEPTLVNGDVAAITMAGEMPMSLGLVDAQAGYQGTALP 237
Db      178 AAKRGGAQVFLGATIGMTTTPGL-----GGATKCFDIDGLVHAQTNVQYQVIF 226
Qy      228 VLVVSMILATIEKFLHKLRLMTADFLIPVLTLLTGLFTFLAIGPAMRWVDLLAHGLO 297
Db      227 VLVAVWVLSTIFKRYHKKLPSAVDTFTPLLSIMITGLFTIIGPVKQVSDAITNIV 286
Qy      298 GLYDFGPGVGLLFGLVSPVIVITGLHOSFPPIELELF-----NQGSGFIETATSMAN 350
Db      287 WLYDTTGAFAQMGVFGISTYSATVTTGLHOSFPFAVETQLAAFAKNPSSGDFIFVACMAN 346
Qy      351 INOGAACLAFLAFLASEKIKGLAGASGSAVAGITEPAIFGVNLTLMRPYIGITGTAIG 410
Db      347 VAQGAATEFVFLTKNKKKKGLASSGVSAALGITTEPALFGVNLKXKPEFFCALIGSGVA 406
Qy      411 GALLALFDIKAVALGAAGFLGVSIDAPDMVFLCAVVTVPVIAFGAIAVGLYLRNNG 470
Db      407 AAFAGIMHTAALASAGLFLSLIPRTIPMWVSVASISFVSLFVVG-----KS 460
Qy      471 SIDPDATAVPAG-----TTQAEAPAEFNSDSTITQAPLTGEALIASVSDAMF 522
Db      461 HKKEVAEODVDVNDATDYAESTQAKELGKQQLKKEIISVPDGPESLTKVNDVDF 520
Qy      522 ASGKIGSVAIVPTKQOLVSPSGKIVVAFPSGHAFAVTKADNSVDILMHIGPTN 582
Db      521 SAKLMGDGAIVPSDSTIYAPVTGYTTIAYETKHAIGI--KSDDA--EVLHIGLDIVN 576
Qy      583 LNGTHENPLKQODEYKAGELLCEPDIDAIKAGAEVTTPIVSN 627
Db      577 LKGEHFESFVKQGRVEKGDKLGSVDLDVAKKAGVDTTVMVVTN 621

RESULT 7
Q74H18 PRELIMINARY; PRT; 647 AA.
AC 074H18;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
RT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DB Phosphoenolpyruvate-dependent sugar phosphotransferase system EIIABC,
   sucrose specific.
GN OrderedlocusNames=LJ0519;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
   Lactobacillus.
ON NCBI_TaxId=33959;
RX NCBI_TaxId=33959;
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Primmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
   Pitter A.-C., Zwalten M.-C., Rouvet M., Altemann E., Barrangou R.,
   Mollet B., Mercenier A., Kleenhammer T., Arigoni F., Schell M.A.;
   "The genome sequence of the probiotic intestinal bacterium
   Lactobacillus johnsoni NCC 533.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517 (2004).
RL EMBL, A5017206; AAS09702.1; -.
DR HSSP, P20166; IAX3.
DR GO, GO:0016020; C:membrane; IEA.
DR GO, GO:0008982; F:protein-N(Pi)-phosphohistidine-sugar phospho. .; IEA.
DR GO, GO:0005351; F:sugar porter activity; IEA.

```

```

DR GO, GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO, GO:0006810; P:transport; IEA.
DR InterPro; IPR01055; Dup_hybrid_motif.
DR InterPro; IPR01535; E1B_glc.
DR InterPro; IPR001996; Pirans_E1B.
DR InterPro; IPR003352; Pirans_E1IC.
DR InterPro; IPR001127; PTS_E1A.
DR InterPro; IPR010973; PTS_T1BC_sucr.
DR Pfam; PF00358; PTS_E1A_1; 1.
DR Pfam; PF00367; PTS_E1B; 1.
DR Pfam; PF02378; PTS_E1IC; 1.
DR ProDom; PD001476; Pirans_E1B; 1.
DR ProDom; PD002243; PTS_E1A; 1.
DR TIGRFAMs; TIGR00826; E1B_glc; 1.
DR TIGRFAMs; TIGR00830; PTBA; 1.
DR TIGRFAMs; TIGR00836; PTS-11-BC_sucr; 1.
DR PROSITE; PS00371; PTS_E1A_1; 1.
DR PROSITE; PS01035; PTS_E1B_CYS; 1.
KW Complete proteome; Pyruvate; Transference.
SQ SEQUENCE 647 AA; 69202 MW; 95FA74C533CE9541 CRC64;

```

```

Query Match 43.3%; Score 1446.5; DB 2; Length 647;
Best Local Similarity 45.0%; Pred. No. 5,86-79;
Matches 300; Conservative 125; Mismatches 209; Indels 33; Gaps 9;

```

```

Qy      1 MDHDLAKRIIRDGEGDNIVAAAHCAITRLVLTDKDVRQSLDPPDKGFEETGGM 60
Db      1 MDHKVAERVAEVA--GRNLTAAAHCAITRLVLTDKDSKNMKNLSDPDKGFKTNGQ 59
Qy      61 FOIIVGPDVHVFEKELDDATS--KQIAVSTBQLKDVANNA--NMFSRAVKVLADIIVPL 118
Db      60 YQVILIGPDVNVVDEFTIKITGLS--ELSTDLDKVAQKQKQFNVMAFIKILSDIFPI 118
Qy      119 PIVVGGGLMANNVLAODLFGPQSLVEMPPQISGVAEMINIMASAPFALPVLVGFTA 178
Db      119 PALVAGGLMALNNFLTSPGLFGKSVYQMAPNVAIGSEIIVQMSAPFIEMPIVVGSA 178
Qy      179 TKRFRGNEFLGAGIGMAMVPEPTLVNGDVAAITMAGEMPMSLGLVDAQAGYQGTALP 238
Db      179 AKRFRANFLDATTIGMTTTPAL-----GGAKTYDIDGLHVSQTNVQYQVIFV 227
Qy      229 LVVSMILATIEKFLHKLRLMTADFLIPVLTLLTGLFTFLAIGPAMRWVDLLAHGLO 298
Db      228 LVAVWLAFLKRPKPKLPSAVDTFTPLLSIMITGLFTIIGPVKQVSDAITNIV 287
Qy      299 LYDFGPGVGLLFGLVSPVIVITGLHOSFPPIELEFNQ-----GSGFIEATSMANIA 352
Db      288 LNTTGARGMGI FGLSYSAIVTTGLHOSFPFAVETQLAEVARGSGDFIFVACMANVA 347
Qy      353 QGAACLAFLAFLASEKIKGLAGASGSAVAGITEPAIFGVNLTLMRPYIGITGTAIGCA 412
Db      348 QGAATFAIYFLTKHEKVGGLASSGVSAALGITTEPALFGVNLKXKPEFFCALIGAGCA 407
Qy      413 LIALFDIKAVALGAAGFLGVSIDAPDMVFLCAVVTVPVIAFGAIAVGLY-----LYR 467
Db      408 FAGLMHTAALASAGLFLSGVSWPKSIPMALSVISFIVAFTFYGRHREKEDVE 467
Qy      468 RNSGIDPDATAVPAGTTKAEAPAEFNSDSTITQAPLTGEALIASVSDAMFASGKL 527
Db      468 ESGTVE--SAGDOVAQGEKAEQIIEKDELHDEVIAPVSGKBSLQDVNDVPFSTAM 524
Qy      528 GSGVAIVPTKQOLVSPSGKIVVAFPSGHAFAVTKADNSVDILMHIGPTNVLNGTH 587
Db      525 GKGAAMVSEGIYSPVGETITVAETGAAGI--KSEGA--EVLHIGLDIVNMKKEG 580
Qy      588 FNPLKQODEYKAGELLCEPDIDAIKAGAEVTTPIVSNYTKTGVPVNTYGLGEIEAGAN 647
Db      581 FTTDVKQGHVEKELGTVLDLAVKAGGYDTTWTITNTTSYANQRIDGVKXGDD 640
Qy      648 LNVAKK 654
Db      641 LIAVTKR 647

```

```

RESULT 8
Q02YRS PRELIMINARY; PRT; 643 AA.
AC Q02YRS;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE PTS system, IABC components.
GN OrderedLocusNames=EFA0067;
OS Enterococcus faecalis (Streptococcus faecalis).
OC plasmid pTEF1.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxId=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.B., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Khouiri H.M., Uettersack T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis."
RL Science 299:2071-2074 (2003).
DR EMBL; AE016833; AAO83060.1; -.
DR HSSP; P20166; 1GPR.
DR TIGR; EFA0067; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup_hybrid_motif.
DR InterPro; IPR001996; Ptrans_EIIB.
DR InterPro; IPR001352; Ptrans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR010973; PTS_II_BC_sucr.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB_1.
DR Pfam; PF02378; PTS_EIIC_1.
DR ProDom; PD001476; Ptrans_EIIB; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMs; TIGR00830; PTBA; 1.
DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
DR Complete proteome.
SQ SEQUENCE 643 AA; 68532 MW; 752P65D8154A06C CRC64;

Query Match 40.3%; Score 1345.5; DB 2; Length 643;
Best Local Similarity 43.1%; Pred. No. 6,9e-73;
Matches 286; Conservative 124; Mismatches 220; Indels 33; Gaps 10;

QY 1 MDHKLQRLIDIGGEDNIVAAHCAETRLVLKDTKVDVROSLDDDDPKLGTPTGGM 60
DB 1 MUYAKLASDIIILAV-GKNDLIAAHGATRLVLKONTVNQKALDENNVNAGTFRIDQ 59
QY 61 FOITVPGDVDFVFEKELDATSKDIAVSTEQLEKDVANNA--NWFSRAVKTLADIVPLI 118
DB 60 YVITIGAGVNFVYDELIKKTGLS-ELSTDDDKQIVDKKKKTPNIPWALIKLSLSEIVPTI 118
QY 119 PLVGGGLMANNVVAQDLFGPOSIVMPQISGVAMINIMASAPAPLPLVVGFPFA 178
DB 119 PLVVGGLMANNVVAQDLFGPOSIVMPQISGVAMINIMASAPAPLPLVVGFPFA 178
QY 179 TTRFGNEFLGAGIGMAWVPTLVNGVDVAATMTAGEMWMSLFGILDVAQAGYQTVLPV 238
DB 179 AARFGANGLGAAIGMIMTTPDL-----GGATEYWNIFGVHVAQTNYAVQVIV 227

```

```

QY 239 LVSMILATIEKFLKRLMGTFADFLTPVLTLTLGLFLFLIAGPMRWGDLAHLGLOG 298
DB 228 IASVYLSTLEKRYEFKRLPSSIDFTPTPLSLVITIGFLFTVIGPTMLLSNGITDAIWM 287
QY 299 LVDFGAPVGLFLFGLVSPYIVITGLHOSFPPIELEFNQ-----GGSFIFATASMANIA 352
DB 288 LVNAGFLGMGIFGQTYSLIVWTGLHOSFPALEQLLAWMNGIGHGDIFFVASMANYA 347
QY 353 QGAACLAFFLAKSEKLGLAGASVAVLGITTEPAIFGVNLRLMPFYIGITGAAGGA 412
DB 348 QGAATFAIWFILTKNSKTSIASAGLSALGITTEPALFGVNLKYFPFPCALIGSIAAA 407
QY 413 LIALDIAVNLGAGFLGVNSIDAPDMWFLVCVAVTFVIAFGAIAAGVLYVRNNGSI 472
DB 408 ITGLKVAVSLSGAGFLGFSINATSLPEYLLCELISVSFAITFYFG--RTSSSIF 465
QY 473 DDDATAAPVAGTAAEAPAEFNSDTI--IQAPLGEALIASVSDAMPASGLGSG 530
DB 466 AAEALAEQTSVTSSEINQIADNPPTVETIVSPPLAGEITTLGSVNDPVSSSIGKG 525
QY 531 VAIVPTKQGLVSPSGKIVAVPSPGHAPAVRTKABDGSNDILMHIGPTVNLNGTHNP 590
DB 526 IAIKENGMTIYSPVDGIVQVVEFEGHAYDKS---NTGAEILIHVGIDTVSLNGKGF- 580
QY 591 LKKQG--DEYKAGELLCERPDIDAIKAGVEYVTPIVSNYKKTGCVNTYGLGEIAGNL 648
DB 581 -KKVAKOKVKKGVGLGTFDSTVITNSGLDITWIVTNSKDYSEVIPITKNIITEGAL 639
QY 649 LNV 651
DB 640 LTI 642

```

```

RESULT 9
Q03ML1 PRELIMINARY; PRT; 534 AA.
AC Q03ML1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Sucrose-specific enzyme II of the PTS (Fragment).
GN Name=screA;
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxId=1599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21822499; PubMed=11832506;
RA Duder A.M., Chailion S., Hissler U., Stentz R.,
RA Champomier-Verges M.C., Albert C.A., Zagorec M.;
RT "Physical and genetic map of the lactobacillus sakei 23K chromosome.";
RL Microbiology 148:421-431 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Duder A.-M., Chailion S., Zagorec M.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF401046; AAK92528.1; -.
DR HSSP; P45618; 2GPR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup_hybrid_motif.
DR InterPro; IPR003352; Ptrans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF02378; PTS_EIIC_1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMs; TIGR00830; PTBA; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
FT NON_TER 1

```

SQ SEQUENCE 534 AA; 56230 MW; 01804F9DE70C0089 CRC64;

Query Match 37.8%; Score 1264; DB 2; Length 534;
 Best Local Similarity 4.76%; Pred. No. 4.6e-68;
 Matches 265; Conservative 99; Mismatches 155; Indels 38; Gaps 10;

114 FVPLPIPIVGGGLMAINNVVAQDLFGPQSLVEMFPOISGVAMINIMASAPAFPLVL 173
 1 FVPLPIPALTAGLLMAINNVVLTGQGLFGAQSIVQMFPOKGFALIVNMSSAPFPLPL 60

174 VFTPTKRGNEPFGAGIGAMVPTLVNGVDVAATAGAMPMSLFGDLVAQAGQG 233
 61 IAFSATKRGNGPYGAAAGMWLVMPNVLNGVGBESLATGHTVWHVFGALIAQAGQG 120

234 TVLPVLVSMILATIEKFLHKLMTADFLTPVTLTLTGFLTAIGPAMRWGDLLA 293
 121 QVPIPIGAVFLANIEKFPKHLNDAVDPFTPMISITITGTLTIVGPAIRIVNGVT 180

294 HGLQGLYDFGSPVGGLLFGLVYSPVITGTHOSPPPIELF---NCGSFTFATASMA 349
 181 DELVMAVQTLGAVGNGIFGLGSAIVLTGLHOSFPALIEFTLLADIATKGGSFIFVAMA 240

350 NTAQGAACLAVERFLAKSEKLKGLAAGSGVSAVLGITEPAIFGVNLRMPFYIGITAI 409
 241 NTAQGAATRAVFVTKNKKQKSLTTSAGISANMLGITBPAFLKXFPFVIGLIAAGI 300

410 GGALIALFDIKAVALGAGFLGVSIDAPDMVFWCAVTVFVIAFGAIAVGLYVERN 469
 301 SSFTIGLHLVLSVSMGPACIGFIAIAPKSPISPMGAIISFVIAFVGTLYGKKAMK-- 358

470 GSIDPATAAPVPAQTAKA--EAPA-----EFSNDSTIIQAPLTGELIALSSVDMF 522
 359 -----TTBEEIINEAPATREVERLUDEK-ISAFTGRVLDIASVPDPVF 402

523 ASGKLGSAIVPTKGLVSPVSGKIVAFPSGHAFAVTRKADSGSNVDILNIGFEDVN 582
 403 ASEMAKGLAIMPDSQDVAIPVGTITIAANTGHANGI--KSDDGA--EVLHIGIDTN 458

583 LNGTHENPLKKQDEYKAGELLCEPDIDAIKAAGYEVTTPIVSN--YKKTGPVNTYGL 639
 459 LNGIGREKIVQCGQHVSEBGLLGHFDIDIKIQAQGLPTLMTITVWAGVADPLTLVDK 518

640 GEIEGAGNLLNV-AKKE 655
 519 AAMQ-GEETIQHAKKD 534

RESULT 10
 PTB_STAXY STANDARD; PRT; 480 AA.

AC P51184;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE PTS system, sucrose-specific IIBC component (EIIIBC-scr) (Sucrose-
 permease IIBC component) (Phosphotransferase enzyme II, BC component)
 DE (EC 2.7.1.69) (EII-Scf).
 GN Name=scrA;
 OS Staphylococcus xylosum.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxId=1288;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 20267 / Isolate C2A;
 RA MEDLINE=94049686; PubMed=8232209;
 RX Wagner E., Goetz F., Bruckner R.,
 RT "Cloning and characterization of the scrA gene encoding the sucrose-
 specific Enzyme II of the phosphotransferase system from
 Staphylococcus xylosum."
 RT Mol. Gen. Genet. 241:33-41(1993).
 RL -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
 CC sugar phosphotransferase system (PTS), a major carbohydrate active
 CC -transport system. The IICD domains contain the sugar binding site

CC and the transmembrane channel; the IIA domain contains the primary
 CC phosphorylation site (the donor is phospho-HPr); IIA transfers its
 CC phosphoryl group to the IIB domain which finally transfers it to
 CC the sugar.
 CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 CC histidine + sugar phosphate.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: Contains 1 PTS EIIb domain.
 CC -1- SIMILARITY: Contains 1 PTS EIIc domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL, X69800; CAA49461.1; -.
 DR PIR, S39978; S39978.
 DR InterPro, IPR001996; Pfam_EIIB.
 DR InterPro, IPR003352; Pfam_EIIC.
 DR InterPro, IPR010973; PTS_II_BC_sucr.
 DR Pfam, PF00367; PTS_EIIB_1.
 DR Pfam, PF02378; PTS_EIIC_1.
 DR ProDom, PD001476; Pfam_EIIB_1.
 DR TIGRFAMs, TIGR00826; EIIb_glc_1.
 DR TIGRFAMs, TIGR01995; PTS-II-ABC-Delta_1.
 DR TIGRFAMs, TIGR01996; PTS-II-BC-sucr_1.
 DR TIGRFAMs, TIGR01992; PTS-II-BC-Tre_1.
 DR PROSITE, PS01035; PTS_EIIB_CYS_1.
 KM Inner membrane; Phosphorylation; Phosphotransferase system;
 KM Sugar transport; Transferase; Transmembrane.
 FT DOMAIN 1
 FT MOD RES ? 480 EIIb.
 FT MOD RES 26 26 Phosphocysteine (By similarity).
 FT MOD RES 325 325 Phosphohistidine (By similarity).
 SQ SEQUENCE 480 AA; 51326 MW; AB4E1D9785D84E47 CRC64;

Query Match 36.7%; Score 1227; DB 1; Length 480;
 Best Local Similarity 52.7%; Pred. No. 7e-66;
 Matches 252; Conservative 83; Mismatches 133; Indels 10; Gaps 6;

1 MDHKLARITLRDGGENIVAAACATRLTLVKDTCDVDRQSLDDPDLKFTETGGM 60
 1 MMYKASANIILQALGGEENVEMTHCATRLTLVKDBGVBEKALGDMVDVKGTFSTGG 60

61 FOITVPGDVHVFKEKLDATSKDIAVSTEQDKDVANNANWFSRAVKLTADIFVPLIP 120
 61 YQVITGSGTVKVFSELEKITGKE-ASSVSEVKTQGTGNMPPORFVWGLSDIFVPLIPA 119

121 LVGGGLMAINNVVAQDLF-GPQSLVEMFPOISGVAMINIMASAPAFPLVAVGFAT 179
 120 IVAGGLMGININILTPAIGFYDNQSLIEVQFSGLAEMINIFNAPFTLLPILIGSAA 179

180 KRFGNIEPFGAGIGAMVPTLVNGVDVAATAG-EMPMMSLFGDLVAQAGQGVTLV 238
 180 KRFGNALVAGLGIILVHPBELMSAYDYPKXLEKKEIPHNNGLEINQVGGQGVLP 239

239 LVVSMILATIEKFLHKLMTADFLTPVTLTLTGFLTAIGPAMRWGDLLAHGQ 298
 240 LVATVTLATIEKGLAKVPTVLNLTPLLAIIISGFTFSFVGTLTTLGYWISDGLTW 299

299 LVYRGPGVGLLFGVYSPVITGTHOSPPPIELF---NCGSFTFATASMANIAQG 354
 300 IVEFGAALGILFGLVAPVITGTHSHFALIEVTLIDSSSTGSGFTFPAITMSNIQG 359

355 AACLAVERFLAK-SEKLKGLAAGSGVSAVLGITEPAIFGVNLRMPFYIGITAIAGAL 413
 360 AAALAAFPITKENKTLKGVASAGSALGITEPMPFVNLRMPFGALVSGISGAY 419

414 IALFDIKAVAGAGFLGVSIDADP--MVMFLCAVTVFVIAFGAIAVGLYVERN 469


```

Db      420  IAFKVKVXALAGTACIIPGISTISGQNNGLHGMIIIAFYAFGVYALSTRKXYRN 477

RESULT 11
O8CN82 PRELIMINARY; PRT; 481 AA.

ID      O8CN82
AC      O8CN82
DT      01-MAR-2003 (TREMBLrel. 23, Created)
DT      01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT      01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE      PTS system sucrose-specific IIBC component.
GN      OrderedLocusNames=SEI1959;
OS      Staphylococcus epidermidis.
OC      Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX      NCBI_taxId=1282;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 12228;
RX      PubMed=12395922;
RA      Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA      Qian Z.-Q., Mao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA      Yun Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT      "Genome-based analysis of virulence genes in a non-biofilm-forming
RT      Staphylococcus epidermidis strain (ATCC 12228).";
RL      Mol. Microbiol. 49:1577-1593(2003).
DR      EMBL; AE016750; AAC05600.1; -.
DR      HSBP; P05053; IIBA.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
DR      GO; GO:0005351; F:sugar porter activity; IEA.
DR      GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR      GO; GO:0006810; P:transport; IEA.
DR      InterPro; IPR011535; E1IB_glc.
DR      InterPro; IPR011966; Ptrans_E1IB.
DR      InterPro; IPR003352; Ptrans_E1IB.
DR      InterPro; IPR010973; PTS_II_BC_sucr.
DR      Pfam; PF00367; PTS_E1IBc_1.
DR      Pfam; PF03378; PTS_E1IBc_1.
DR      ProDom; PD001476; Ptrans_E1IB; 1.
DR      TIGRFAMs; TIGR00826; E1IB_glc; 1.
DR      TIGRFAMs; TIGR01966; PTS-II-BC-sucr; 1.
DR      PROSITE; PS01035; PTS_E1IB_Cys; 1.
KW      Complete proteome.
SQ      SEQUENCE 481 AA; 51609 MW; 447F2BE3B04F409 CRC64;

Query Match      35.5%; Score 1186; DB 2; Length 481;
Best Local Similarity 49.9%; Pred. No. 2,1e-63;
Matches 235; Conservative 95; Mismatches 123; Indels 18; Gaps 6;

Oy      1 MDHKLARILRDIGEDNIVAAACARLRVLKDTDVDRQSLDDDDDLKGTETEGM 60
Db      1 MSYKSAEELILVAIGSEENLDMACARLRVLNDESKVEDDTLSNMDVNGITSTGQ 60
Oy      61 FOIIVSPDVDVHVFEGELDATSKDIAVSTEQLDVYANNANMFSAVKVLADI FPLPI 120
Db      61 VOIIVSGSTVKNVFELEKIKTKE-ASITSEVKDSSKGMNPFQCFVKMLSTIFPIIPA 119
Oy      121 LVGGGLMAINNVLAODLF-GPQSLVENFPOISGVAENINIMASAPPAFLPVLVGPTAT 179
Db      120 IVAGGLLGLMGLNNIIFAKDLFYDGSKIIVHVSQFSGGLADMINIFANPFLTLPILIGSAA 179
Oy      180 KRFGNENFLGAGIGAMVPPTLVNVDVVAATMTAGE-HPMWSLFGLDVAAQGYOQTVLPV 238
Db      180 KRFGNENFLGAGIGAMVPPTLVNVDVVAATMTAGE-HPMWSLFGLDVAAQGYOQTVLPV 239
Oy      239 LVVSNILATIEKFLHKLMTGADPFLITPVLTLLLTGFLTFIAIGPAMRWGDLAHGLQ 298
Db      240 LVATYIILATIEKFLKRVIPVLDNLTLPLLSIFIAFTIFPLVGVTVRQLGWLSDGLTW 299
Oy      299 LYDEGCPVGGLLFGLVSPYIVTGLHQSFPPELELF---NQGSSTPATASNANIAG 354
Db      300 LYDEGATGGLIFGLLVAPIVTGMHSFIAETLLIDATGTGSLFPPIATMGNIAQ 359

```

```

Oy 355 AACCAVAFPLAK-SETKKRLACAGSVAVLGTEPRLPGVNLRLRNPFIYIGTAAIGCL 413
Db 360 GAALAAFPRIIKONKKLKGVASAGISALIGITEPAMFGVNLRLRPFIGAVAGSGIAY 419
Oy 414 IALPIKRVALLGAAGFLGVSTIDADD-----MMVFLCAVTFVIA 454
Db 420 ISPFKVAIALGTAGLPGFISINPHAGMLHYLIGMLIAFVSVVTLVLVS 470

RESULT 12
O8NV35
O8NV35 PRELIMINARY; PRT: 480 AA.
AC O8NV35;
DT 01-OCT-2002 (TremBRel. 22, Created)
DT 01-OCT-2002 (TremBRel. 22, Last sequence update)
DE 01-MAR-2004 (TremBRel. 26, Last annotation update)
DE PTS system sucrose-specific IIBC component.
GN Name=ecrA; OrderedLocustNames=MM2299;
OS Staphylococcus aureus (Strain MM2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Naai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hizamatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827 (2002).
DR EMBL; AP004830; BAB96164.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011535; E1IB_glc.
DR InterPro; IPR001996; Pctans_E1IB.
DR InterPro; IPR003352; Pctans_E1IC.
DR InterPro; IPR010973; PTS_TT_BC_sucr.
DR Pfam; PF00367; PTS_E1IB_1.
DR Pfam; PF02378; PTS_E1IC_1.
DR ProDom; PD001476; Pctans_E1IB_1.
DR TIGRFAMs; TIGR00826; E1IB_glc_1.
DR TIGRFAMs; TIGR01996; PTS-TT-BC-sucr_1.
DR PROSITE; PS01035; PTS_E1IB_Cys_1.
KW Complete proteome.
SQ SEQUENCE 480 AA; 51218 MW; B3EBE81C08C0DEDA CRC64;

Query Match 35.4%; Score 1183; DB 2; Length 480;
Best Local Similarity 50.4%; Pred. No.3,1e-63;
Matches 242; Conservative 94; Mismatches 130; Indels 14; Gaps 7

Oy 1 MDHKDLQRIIRDIGGEDNIVAAAHACARLLVLKTDKDVDRQSLDDPDLKGTETGGM 60
Db 1 MMYQSAEILINAIIGGENLDMAHCAIRLLVLNDBESVLNEEALNMNDVYKGTSTGCG 60
Oy 61 FOIIVPGDVVDFREKLDATSKDIAVTEQDKDVANNANMFSAVKVLADIYVPLPI 120
Db 61 YQIIGSGTVKVFSELEKLTGKE-ASTTSEKQASAKMNPDLQRFVKCLSDIFVPIIPA 119
Oy 121 LVGGGLMAINNVVVAOOL-FGPGSIVMPFOISGVAEMINLMAAPAFVLVGFNT 179
Db 120 IVAAGGLMGLNNLITAKDLFFSGSLDIDVYQFAGLAEMINVFANAAPFTLLTIGFSA 179
Oy 180 KRFGGNEFLGAGIGMAWFPFLVNGYDVAAATMTAGE-MPMNSFLGLDVAQAGYQGTVPV 238
Db 180 KRFGGNPFLGALGALMIIVHPSLMSAYDPRKAVEGAKAIPYVDVDFGLHINQGYQGQVLP 239
Oy 239 LVVSMILATTEKFLHAKRLMGADPLITPVLTLTLTGFLTPFAIGPMKRVGDLAHLQG 298
Db 240 LVAVYIILASIKGRVKTPTVLDNLTLTFLSLIFITAFVTFVSGITRQLGWTSLDGLTW 299

```

```

QY 299 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFPTASMANIAOG 354
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFPTASMANIAOG 359
QY 335 AACLAFFFLAK-SEKTKGLAGASGVSAVLGITBPAIFGVNLRMPFYIGTAAIGAL 413
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 GAATAAFPIIKONKKLKGVASAGISALLGITBPAIFGVNLRMPFYIGTAAIGAL 419
QY 414 IALFDIKVAALGAGLGVASIDA--PDWVFLVCAVTVFYIAGAAIAYGLYVRNGS 471
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 IAFKVAALGAGLGVASIDA--PDWVFLVCAVTVFYIAGAAIAYGLYVRNGS 475

```

RESULT 13

06G6U2 PRELIMINARY; PRT; 480 AA.

```

ID 06G6U2
AC 06G6U2
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE PTS system, sucrose-specific IIBC component (EC 2.7.1.69).
GN OrderedLocustNames=SA52269;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;

```

SEQUENCE FROM N.A.

```

RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Baason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagsis K.,
RA James K.D., Leonard N., Line A., Mayes R., Moulé S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch B., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571857; CAG4082.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR011535; E1IB_glc.
DR InterPro: IPR001996; Ptrans_E1IB.
DR InterPro: IPR003352; Ptrans_E1IC.
DR InterPro: IPR010973; PTS_II_BC_sucr.
DR Pfam: PF00367; PTS_E1IB; 1.
DR Pfam: PF02378; PTS_E1IC; 1.
DR ProDom: PD001476; Ptrans_E1IB; 1.
DR TIGRFAMs: TIGR00826; E1IB_glc; 1.
DR TIGRFAMs: TIGR01996; PTS_II_BC_sucr; 1.
DR PROSITE: PS01035; PTS_E1IB_Cys; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 480 AA; 51218 MW; B3EBE81C08CODEDA CRC64;

```

Query Match

Best Local Similarity 35.4%; Score 1183; DB 2; Length 480;

Matches 242; Conservative 94; Mismatches 130; Indels 14; Gaps 7;

```

QY 1 MDHKLADGRLRDIGSDENIVAAACATRLRLVKDTKVDVRSQSLDDPDLKGTETGGM 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MMYKGAEEIINAIGEENLDMAHQATRLRLVINDESLVNEEALNNMVGCTSTGGQ 60
QY 61 FOIIVGPGVDVHFKELDQATSKDIAVSTEQKQDVANNANMFSAVKYLAIDFVPLPI 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YOIIGSGTVNKVSELEKLTGKE-ASTTSEVVAQASAKMNPQRFVKVLSDFVPIIPA 119

```

```

QY 121 LVGGGLMANNVIVADQL-FGPOSIVMEPQISGVAEMINIMASAPAPLPVUGFTAT 179
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 IVAQGLMGLNNILTLADDFESGSLIDVYSGFGLAEMINIVFNAPFTLLPILIGSAA 179
QY 180 KRFGNEPLGAGIGAMVFPFLTVNGVDVATMTAGE-MPMWSLSDQVDAQYOGTVLPV 238
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 KRFGNEPLGAGIGAMVFPFLTVNGVDVATMTAGE-MPMWSLSDQVDAQYOGTVLPV 239
QY 239 LVSMILATIEKFLHKLGMGTADLPITPVTLTLITGLVTLAIGPAMRWGDLAHLQOG 298
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 LVAAYIILASIEKGRKXVLPVLDMLTLPLSLPITAFLETFSFVGPITRQGLYMSDGLTW 299
QY 299 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFPTASMANIAOG 354
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFPTASMANIAOG 359
QY 335 AACLAFFFLAK-SEKTKGLAGASGVSAVLGITBPAIFGVNLRMPFYIGTAAIGAL 413
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 GAATAAFPIIKONKKLKGVASAGISALLGITBPAIFGVNLRMPFYIGTAAIGAL 419
QY 414 IALFDIKVAALGAGLGVASIDA--PDWVFLVCAVTVFYIAGAAIAYGLYVRNGS 471
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 IAFKVAALGAGLGVASIDA--PDWVFLVCAVTVFYIAGAAIAYGLYVRNGS 475

```

RESULT 14

06G6S4 PRELIMINARY; PRT; 480 AA.

```

ID 06G6S4
AC 06G6S4
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE PTS system, sucrose-specific IIBC component (EC 2.7.1.69).
GN Name=StcIA; OrderedLocustNames=MSA2466;
OS Staphylococcus aureus (strain MSA2466).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;

```

SEQUENCE FROM N.A.

```

RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Baason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagsis K.,
RA James K.D., Leonard N., Line A., Mayes R., Moulé S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch B., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571856; CAG41448.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR011535; E1IB_glc.
DR InterPro: IPR001996; Ptrans_E1IB.
DR InterPro: IPR003352; Ptrans_E1IC.
DR InterPro: IPR010973; PTS_II_BC_sucr.
DR Pfam: PF00367; PTS_E1IB; 1.
DR Pfam: PF02378; PTS_E1IC; 1.
DR ProDom: PD001476; Ptrans_E1IB; 1.
DR TIGRFAMs: TIGR00826; E1IB_glc; 1.
DR TIGRFAMs: TIGR01996; PTS_II_BC_sucr; 1.
DR PROSITE: PS01035; PTS_E1IB_Cys; 1.
KW Complete proteome.
SQ SEQUENCE 480 AA; 51176 MW; EDF9B97031C26A12 CRC64;

```

Query Match

Best Local Similarity 35.3%; Score 1179; DB 2; Length 480;

	Matches	241; Conservative	95; Mismatches	130; Indels	14; Gaps	7
QY	1	MDHKLQARILIDIGGEDIVAAAHCAATRLRLVKTQKVDKQSLDDDDDKLKEFTGGM	60			
Db	1	MAVKQSAIEDIILNAGEENLDMAHCAATRLRLVLANDESLVNEILANMDVVKKTFSTGQ	60			
QY	61	FOIIVPGDVDVHFKELDDATSKDILAVSTEQKLDVVAANNANMFSAVKYLADIFFVLIPI	120			
Db	61	YQIITIGSGVNVKVFSELEKLTGKE-ASTTSEVKAQSAKRMNPQGRVVKMLSDIFVPIIA	119			
QY	121	LVGGGLMIAINNVVAQDL-FEPQSLVEKMFPOISGVAEINIMLASAPFAFLPVLVGFAT	179			
Db	120	IVAGGLMLGNLNIILTAQDLFFSGKSLIDVYSQFAGLAEMINVFANPFLPILITGFSAA	179			
QY	180	KRFGGNEFGAGIGMANVPEPTLVNGCDVAVATATAGE-MEMMSLFGLDVAQAQGVTLPV	238			
Db	180	KRFGGNPFGLGALGNLILVHPSLSMAVDPEKVAEGAKAIPLYMDVFGHINQVQGVLEH	239			
QY	239	LVVSIILATIEKFLHKRLMGADFLITPVLTLLTGELTFPIAIGPMRVGDLIAHGLQG	298			
Db	240	LVAAYILASIEGKKRKVLPVLDNLTLPILSFITAFLETFSEFGPTROLGWLISGLTW	299			
QY	299	LYDFGAPVGLLFGKLVSPYIVITGLHSPFPIELEF-----NQGSSEFIPATSMANIAQG	354			
Db	300	LYEFGALIGGLIFGLILXAPIYITGMHSHIAVETTLIAATKTGGSFIPPIATMSVAAQG	359			
QY	355	AACLAIVFLLAK-SEKIKGLIAGASGVSAVIGITEPALFGVNLRLRMWFIYIGITAAIGAL	413			
Db	360	GAALIAAFPIKONKKLKGVASAAGISALGITIEPAMFGVNLKLRYPFIGAIVQSGIGSAY	419			
QY	414	IALPIKIKVALGAAGELGVASIDA--PDWVMELVCAVVFVYIAFGAIAVGLYVLRNRS	471			
Db	420	IAFVKKAIALGTAGLPGFTSINPVHAGMLHIFVGKTIISFIV-----AIVTVLLISRRKAN	475			

RESULT 15	099RQ0	PRELIMINARY;	PRT;	480 AA.
ID	099RQ0			
AC	099RQ0			
DT	01-JUN-2001 (TREMBlrel. 17, Created)			
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)			
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)			
DE	PTS system, sucrose-specific IIBC component.			
GN	Name=acra; OrderedLocustNames=SAV3377;			
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699).			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=158878;			
RY	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Mu50 / ATCC 700699;			
RX	MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;			
RA	Kuroda M., Ohta T., Uchihyama I., Baba T., Yuzawa H., Kobayashi I.,			
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,			
RA	Kanamori M., Maeumaru H., Maruyama A., Murakami H., Hosooyama A.,			
RA	Matsumi-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,			
RA	Sekimizu K., Hatakeya H., Kubura S., Goto S., Yabuzaki J.,			
RA	Kanehisa M., Yamaehita A., Oshima K., Furuya K., Yoshino C., Shiba T.,			
RT	Hattori M., Ogasawara N., Hayashi K., Hiratazu K.;			
RI	"Whole genome sequencing of methicillin-resistant Staphylococcus			
RL	aureus.";			
RL	Lancet 357:1225-1240(2001).			
DR	EMBL; AP003365; BAB58539.1; -.			
DR	PIR; D90038; D90038.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0008992; F:phosphatidylserine-sugar phospho. .; IEA			
DR	GO; GO:0005351; F:sugar porter activity; IEA.			
DR	GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA			
DR	GO; GO:0006810; P:transport; IEA.			
DR	InterPro; IPR011535; E1IB_Glc.			
DR	InterPro; IPR001996; E1IB_E1IB.			
DR	InterPro; IPR003352; Ptrans_E1IC.			
DR	InterPro; IPR010973; PTS_11_Bc_sucr.			
DR	Pfam; PF00367; PTS_E1IB; 1.			

DR Pfam: PF02378; PTS_EIIC_1.
DR ProDom: PD001476; Pfam: E1IB_1.
DR TIGRfam: TIGR00826; E1IB_glc_1.
DR TIGRfam: TIGR01996; PTS_I-BC-sucr; 1.
DR PROSITE: PS01035; PTS_E1IB_Cys_1.
DR Complete proteome.
SQ SEQUENCE 480 AA; 51232 MW; D88607F6E0AF2E84 CRC64;

Query Match	35.2%;	Score 1178;	DB 2;	Length 480;
Best Local Similarity	50.4%;	Pred. No. 6.3e-63;		
Matches 242;	Conservative 94;	Mismatches 130;	Indels 14;	Gaps 7;

```

OY 1 MDHKO LARIRLBDJGGBENITVAALACARLRRLVLDXOTKDVDSQJLDBDDBLKGFTGGA 60
Db 1 MNVQSAEDITLNAJGGEENLDMAHACATRLRLVLNDESLVNEBALNMMDVAGSTFTGGQ 60
OY 61 FOIIVBBDVHVHFKELDDATSKDIVSTBQLKDVVANNAMFSAUVKYLADI PVLPIEI 120
Db 61 YOIIGSTQVKNVSELEKTGKE -ASTTSEKQASAKRMNLDORFVKMLSDI PVDI IBA 119
OY 121 LVGGGLMAINNVVAOOL -FGOSLVEMFPDIOGVAEINIMASAPFAFLPVLVGFRTAT 179
Db 120 IVAGGLMGJLNNITLAKDLFPSSGSKLIDVYSOPAGLAEKINPANNAPFLRLPIIGFSA 179
OY 180 KRFGGNEFLGAGIGMAWVPTLVNGVDVAATWTJAGE -MPMWSLFGJLDVAQAGYOGTVLPU 238
Db 180 KRFGGNPFLGAVLGNILVHPSLMSAYDPPKAVEAGKALPYMDVFGHINQVYQOGQVLEPM 239
OY 239 LVWSMLATITEKFLHKKRLMGTADEFILTVLTLITGLFTLFIAGPMRWGDLIAHGLQ 298
Db 240 LVAAYIILASIEKGRKXVPTVLDNLTLLTSLFIATPFLFSPVFGPTRLQGWLSGLTW 299
OY 299 LYDFRGPVGGLFGLVNPSVIVITGSHOSFPPIELEF----NQGSFIPATASMANIAQ 354
Db 300 LYREGOAGGILFGJLVNPIVITGNHHSFIAVETLLIDATITGGSFIPPIATMSVAAQ 359
OY 355 AACLAFFELAK -SEKTKGLAGASGVSAVLGITTEPALFGVNILRLMFPYIGITPAAGAL 413
Db 360 GAALIAAFPIIKONKKLKGVAASAAGISALLGITTEPAMFGVNILKLRVPFIGAIVOSGISAY 419
OY 414 IALDIKRVNLAGAGLGVGSIDA--PDMWFVCAVVTFVAFAGAIYGLVYVRNS 471
Db 420 IAFKRVKAIAGTGLGPGFISINPVHAGMLHFEVMTSFI-----ALVTVLILSRKAN 475

```

Search completed: March 7, 2005, 22:14:08
Job time : 183 secs

THIS PAGE BLANK (uspto)